



MÁSTER UNIVERSITARIO EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER HERRAMIENTA DE PREDICCIÓN DE AFLUENCIA Y OPTIMIZACIÓN DE LA PLANIFICACIÓN DE PERSONAL EN EL SECTOR DE RESTAURACIÓN

Autor: González Molina, Alejandra

Director: Cruz Sánchez de Rojas, Álvaro De La

Coordinador: Ortiz Marcos, Susana

Madrid

Junio de 2019

AUTORIZACIÓN PARA LA DIGITALIZACIÓN, DEPÓSITO Y DIVULGACIÓN EN RED DE PROYECTOS FIN DE GRADO, FIN DE MÁSTER, TESIS O MEMORIAS DE BACHILLERATO

1º. Declaración de la autoría y acreditación de la misma.

El autor D. Alejandra González Molina DECLARA ser el titular de los derechos de propiedad intelectual de la obra: HERRAMIENTA DE PREDICCIÓN DE AFLUENCIA Y OPTIMIZACIÓN DE LA PLANIFICACIÓN DE PERSONAL EN EL SECTOR DE RESTAURACIÓN, que ésta es una obra original, y que ostenta la condición de autor en el sentido que otorga la Ley de Propiedad Intelectual.

2º. Objeto y fines de la cesión.

Con el fin de dar la máxima difusión a la obra citada a través del Repositorio institucional de la Universidad, el autor **CEDE** a la Universidad Pontificia Comillas, de forma gratuita y no exclusiva, por el máximo plazo legal y con ámbito universal, los derechos de digitalización, de archivo, de reproducción, de distribución y de comunicación pública, incluido el derecho de puesta a disposición electrónica, tal y como se describen en la Ley de Propiedad Intelectual. El derecho de transformación se cede a los únicos efectos de lo dispuesto en la letra a) del apartado siguiente.

3º. Condiciones de la cesión y acceso

Sin perjuicio de la titularidad de la obra, que sigue correspondiendo a su autor, la cesión de derechos contemplada en esta licencia habilita para:

- a) Transformarla con el fin de adaptarla a cualquier tecnología que permita incorporarla a internet y hacerla accesible; incorporar metadatos para realizar el registro de la obra e incorporar “marcas de agua” o cualquier otro sistema de seguridad o de protección.
- b) Reproducirla en un soporte digital para su incorporación a una base de datos electrónica, incluyendo el derecho de reproducir y almacenar la obra en servidores, a los efectos de garantizar su seguridad, conservación y preservar el formato.
- c) Comunicarla, por defecto, a través de un archivo institucional abierto, accesible de modo libre y gratuito a través de internet.
- d) Cualquier otra forma de acceso (restringido, embargado, cerrado) deberá solicitarse expresamente y obedecer a causas justificadas.
- e) Asignar por defecto a estos trabajos una licencia Creative Commons.
- f) Asignar por defecto a estos trabajos un HANDLE (URL *persistente*).

4º. Derechos del autor.

El autor, en tanto que titular de una obra tiene derecho a:

- a) Que la Universidad identifique claramente su nombre como autor de la misma
- b) Comunicar y dar publicidad a la obra en la versión que ceda y en otras posteriores a través de cualquier medio.
- c) Solicitar la retirada de la obra del repositorio por causa justificada.
- d) Recibir notificación fehaciente de cualquier reclamación que puedan formular terceras personas en relación con la obra y, en particular, de reclamaciones relativas a los derechos de propiedad intelectual sobre ella.

5º. Deberes del autor.

El autor se compromete a:

- a) Garantizar que el compromiso que adquiere mediante el presente escrito no infringe ningún derecho de terceros, ya sean de propiedad industrial, intelectual o cualquier otro.
- b) Garantizar que el contenido de las obras no atenta contra los derechos al honor, a la intimidad y a la imagen de terceros.
- c) Asumir toda reclamación o responsabilidad, incluyendo las indemnizaciones por daños, que pudieran ejercitarse contra la Universidad por terceros que vieran infringidos sus derechos e

intereses a causa de la cesión.

- d) Asumir la responsabilidad en el caso de que las instituciones fueran condenadas por infracción de derechos derivada de las obras objeto de la cesión.

6º. Fines y funcionamiento del Repositorio Institucional.

La obra se pondrá a disposición de los usuarios para que hagan de ella un uso justo y respetuoso con los derechos del autor, según lo permitido por la legislación aplicable, y con fines de estudio, investigación, o cualquier otro fin lícito. Con dicha finalidad, la Universidad asume los siguientes deberes y se reserva las siguientes facultades:

- La Universidad informará a los usuarios del archivo sobre los usos permitidos, y no garantiza ni asume responsabilidad alguna por otras formas en que los usuarios hagan un uso posterior de las obras no conforme con la legislación vigente. El uso posterior, más allá de la copia privada, requerirá que se cite la fuente y se reconozca la autoría, que no se obtenga beneficio comercial, y que no se realicen obras derivadas.
- La Universidad no revisará el contenido de las obras, que en todo caso permanecerá bajo la responsabilidad exclusiva del autor y no estará obligada a ejercitar acciones legales en nombre del autor en el supuesto de infracciones a derechos de propiedad intelectual derivados del depósito y archivo de las obras. El autor renuncia a cualquier reclamación frente a la Universidad por las formas no ajustadas a la legislación vigente en que los usuarios hagan uso de las obras.
- La Universidad adoptará las medidas necesarias para la preservación de la obra en un futuro.
- La Universidad se reserva la facultad de retirar la obra, previa notificación al autor, en supuestos suficientemente justificados, o en caso de reclamaciones de terceros.

Madrid, a 20 de Junio de 2019

ACEPTA



Fdo.....

Motivos para solicitar el acceso restringido, cerrado o embargado del trabajo en el Repositorio Institucional:

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título
HERRAMIENTA DE PREDICCIÓN DE AFLUENCIA Y OPTIMIZACIÓN DE LA
PLANIFICACIÓN DE PERSONAL EN EL SECTOR DE RESTAURACIÓN
en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el
curso académico 2018/2019 es de mi autoría, original e inédito y
no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro,
ni total ni parcialmente y la información que ha sido tomada
de otros documentos está debidamente referenciada.



Fdo.: Alejandra González Molina

Fecha: 20 / 06 / 2019

Autorizada la entrega del proyecto
EL DIRECTOR DEL PROYECTO



Fdo.: Álvaro de la Cruz Sánchez de Rojas

Fecha: 20 / 06 / 2019



MÁSTER UNIVERSITARIO EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER HERRAMIENTA DE PREDICCIÓN DE AFLUENCIA Y OPTIMIZACIÓN DE LA PLANIFICACIÓN DE PERSONAL EN EL SECTOR DE RESTAURACIÓN

Autor: González Molina, Alejandra

Director: Cruz Sánchez de Rojas, Álvaro De La

Coordinador: Ortiz Marcos, Susana

Madrid

Junio de 2019

HERRAMIENTA DE PREDICCIÓN DE AFLUENCIA Y OPTIMIZACIÓN DE LA PLANIFICACIÓN DE PERSONAL EN EL SECTOR DE RESTAURACIÓN.

Autor: González Molina, Alejandra

Director: Cruz Sánchez de Rojas, Álvaro D

Entidad Colaboradora: PricewaterhouseCoopers Asesores de Negocio S.L

RESUMEN DEL PROYECTO

1. Introducción

El sector restauración es actualmente uno de los sectores más importantes para la economía española. Además, se trata de un elemento clave en el sector turístico ya que genera un gran atractivo para el mismo.

Tras superar un largo período de crisis y a pesar de haber sido uno de los sectores más castigados durante dicho periodo, en 2017 el valor del mercado registró una evolución positiva por cuarto año consecutivo. Fue en 2014 cuando tras haber tocado fondo después de seis ejercicios consecutivos a la baja, se produjo un cambio de tendencia en el mercado.

Cabe destacar que la fuerte recuperación que se está produciendo en el sector es debida principalmente a un aumento en el tráfico o número de visitas realizadas a los locales de hostelería, y no a un aumento en el ticket medio consumido. [KPMG18] A pesar de ello, el sector de la hostelería en conjunto ha perdido un 14% de facturación desde que estalló la crisis en 2008.

Por otro lado, los hábitos de consumo, los gustos y los comportamientos de los clientes, así como los tipos de experiencias gastronómicas se encuentran en plena evolución. Es por ello que las empresas se están viendo obligadas a adaptarse a este cambio transformando su tradicional modelo de negocio, incorporando las nuevas tecnologías en él. Además, las ofertas cada día están más orientadas a satisfacer a los clientes, que cada vez están más informados, son más exigentes, buscan más innovación y son menos fieles a las marcas. A su vez, el auge de la economía colaborativa y la fuerte penetración de los smartphones están empujando a este sector hacia la digitalización.

Los factores claves en esta transformación que está experimentando el sector de restauración están siendo el uso de herramientas para el análisis de datos, la comunicación directa con los consumidores a través de redes sociales y otras plataformas, y el desarrollo de nuevos servicios que aportan verdadero valor añadido a los clientes [KPMG18] .

Cabe destacar que la planificación es la clave del éxito en los restaurantes. La gestión del personal y la gestión del stock son dos de los aspectos más importantes a tener en cuenta a la hora de lograr buenos resultados. El personal es uno de los activos más importantes de un restaurante ya que el número de ventas dependerá en gran parte de su productividad y del servicio que ofrezcan. Además, la rapidez de servicio y la calidad del mismo se traducen directamente en satisfacción del cliente y en incremento de ventas.

Gestionar de manera óptima los turnos y horarios de un restaurante, así como el número de empleados necesarios para cada momento del día se trata de una labor de gran complejidad que

requiere una gran inversión de tiempo y esfuerzo. A día de hoy, no se le otorga a esta actividad la importancia que realmente tiene ya que en gran parte se desconoce el impacto de la misma sobre los resultados de un restaurante, así como los beneficios que se podrían obtener por una óptima gestión de esta tarea.

Actualmente, en la gran mayoría de restaurantes la herramienta de planificación de personal más utilizada es Excel. Normalmente se utilizan plantillas de turnos generales, en las cuáles se va asignando manualmente el personal necesario y los horarios correspondientes. En general, en estos métodos de planificación no se tienen en cuenta las posibles fluctuaciones que pueden llegar a ocurrir en un corto plazo como consecuencia de factores como el tiempo atmosférico, acontecimientos puntuales, factores sociales, factores económicos... para prever el personal necesario. Esto conlleva situaciones amenazadoras y de grave peligro para los restaurantes como el fenómeno conocido como Under/Over-Staffing, que hace referencia al exceso de personal o la falta del mismo.

Según algunas fuentes, el exceso de horas planificadas en restaurantes puede llegar a superar el 15%. Por otro lado, la falta de personal en ciertos horarios puede provocar pérdidas de aproximadamente el 4%. Cualquiera de estas dos situaciones supone una gran pérdida de beneficios para la empresa, así como una disminución en la satisfacción del cliente y un deterioro de la imagen de marca y la reputación.

A pesar de existir herramientas de organización de personal y creación de turnos, este tipo de herramientas no permite la adaptación de los horarios y turnos de los empleados a las necesidades específicas y previstas del restaurante, y lo más importante, no suelen incorporar variables externas al negocio, que tiene un gran impacto en la afluencia. Es decir, no realiza una previsión analítica de las necesidades que va a tener el propio establecimiento teniendo en cuenta diversos factores para actualizar los turnos y número de empleados necesarios en tiempo real. Por otro lado, dos nuevas herramientas en el mercado afirman utilizar analítica predictiva para estimar la afluencia futura y elaborar la planificación en consecuencia. Sin embargo, existe una gran dificultad a la hora de estandarizar las casuísticas sobre una herramienta de mercado, por lo que es posible que dichas herramientas puedan perder precisión y personalización de cara al cliente, además de incurrir en elevados costes.

2. Metodología

El objetivo principal del proyecto es el desarrollo de una herramienta que permita incorporar la base de datos de un cliente real a la misma y que sea capaz de, a partir de esos datos históricos realizar una predicción de afluencia teniendo en cuenta diversos factores y calcular el número de empleados óptimo para cada turno. La herramienta incorporará algoritmos de aprendizaje supervisado (Machine Learning) e inteligencia artificial para obtener los resultados [COSS18].

El proyecto estará estructurado en tres partes: la simulación del negocio de restauración, la predicción de afluencia para una cadena de restauración simulada y la optimización de la gestión del personal teniendo en cuenta la afluencia prevista para la cadena.

Para la realización de la simulación del negocio se empleará el método Delphi. Este método consiste en reunir a un grupo de expertos con los que se discutirán y validarán de manera iterativa los resultados obtenidos en el proceso de construcción de la simulación del negocio hasta conseguir una simulación que refleje la realidad de la manera más precisa.

Una vez conseguida la simulación del negocio, se empleará como base de datos de la cadena de restauración simulada. A continuación, se procederá a realizar la predicción de afluencia. Para ello se implementarán dos modelos matemáticos de predicción: Random Forest y Prophet.

El objetivo es comparar los resultados y verificar qué modelo ofrece una solución más precisa de predicción de afluencia. Estos modelos matemáticos se desarrollarán, parametrizarán y ejecutarán en el entorno de programación Python.

Finalmente, una vez obtenidos los datos estimados de predicción de afluencia que tendrá la cadena de restauración se procederá a optimizar la planificación de empleados de la empresa. Para ello se realizarán una serie de hipótesis y se establecerán una serie de estándares de calidad para poder maximizar los beneficios ofreciendo el mejor servicio posible.

Con el objetivo de demostrar la eficacia de esta herramienta y el impacto que puede tener en el mercado se elaborará un caso de negocio y se analizarán distintos escenarios ofreciendo una serie de conclusiones al respecto.

3. Resultados

A continuación, se muestra la interfaz de la herramienta de visualización creada que presenta los resultados obtenidos en cada una de las tres partes mencionadas. La primera interfaz muestra el histórico del restaurante junto con los indicadores claves de rendimiento; la segunda incluye los resultados de ambos modelos de predicción para poder comparar y verificar la eficacia y precisión de cada uno; la tercera pantalla refleja el número de empleados óptimo que sería necesario en cada franja horaria del turno para maximizar beneficios junto con la comparativa del modelo de planificación actual. También se podrá observar en esta última pantalla una comparación de los beneficios actuales frente a los que se obtendrían en el caso óptimo.

Las tres pantallas ofrecen la posibilidad de seleccionar una serie de filtros con la que se actualizará la información mostrada, con el objetivo de poder personalizar la información que se desea visualizar y analizar ciertos resultados en particular.

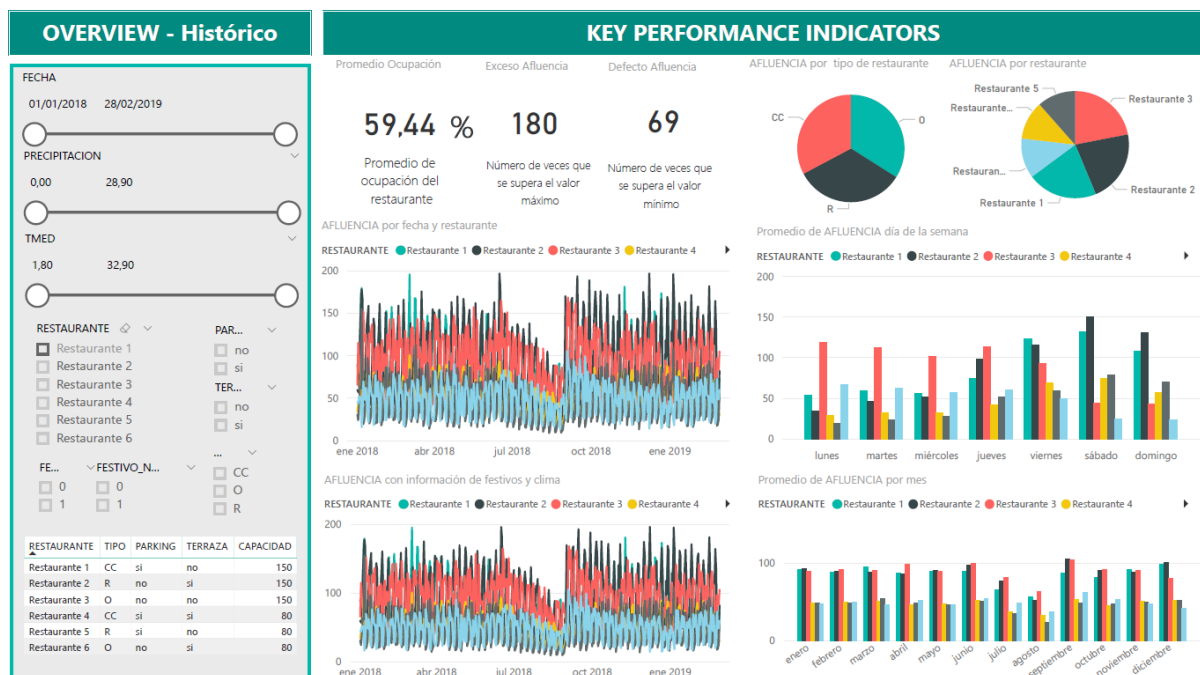


Figura 1. Primera interfaz de la herramienta

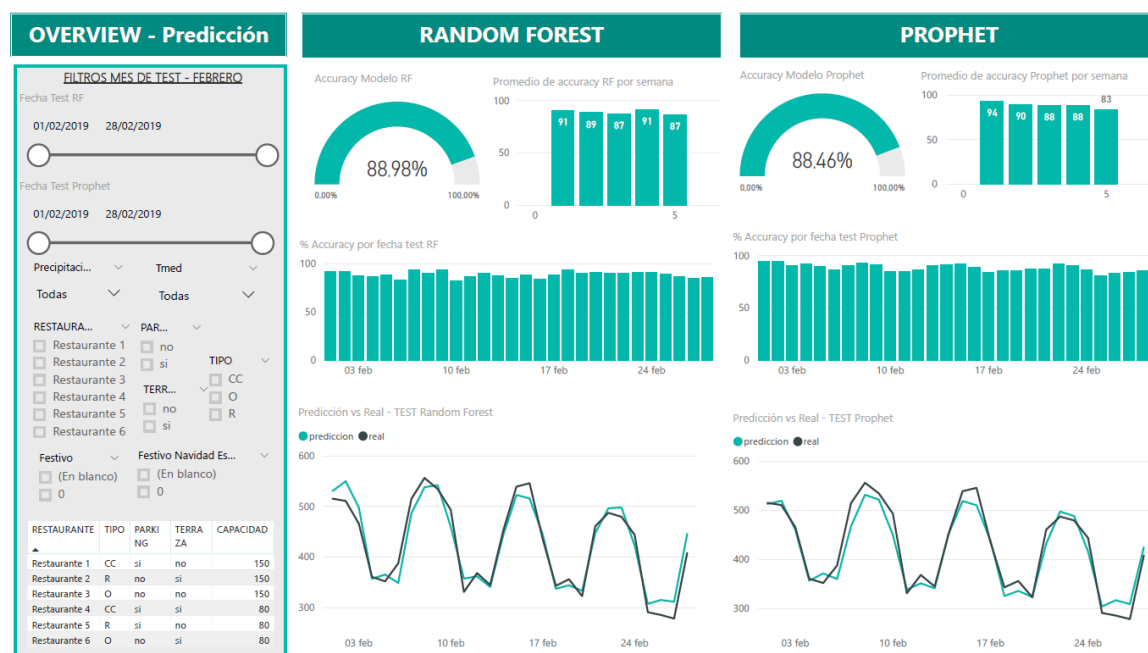


Figura 2. Segunda interfaz de la herramienta

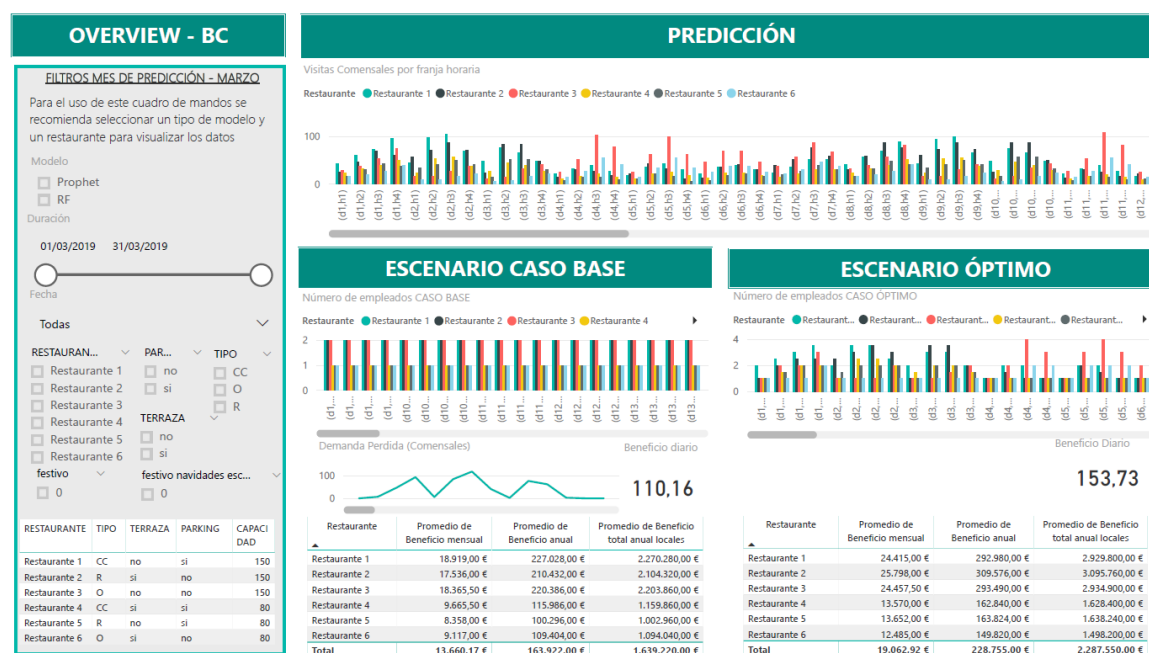


Figura 3. Tercera interfaz de la herramienta

La simulación del negocio fue aceptada como válida tras numerosas iteraciones de corrección y validación con el grupo de expertos, incluyendo diversos factores que tienen influencia directa en la afluencia de comensales en restaurantes.

Respecto a los resultados obtenidos en el apartado de predicción, se ha observado que ambos modelos serían válidos para la predicción de afluencia en este caso puesto que presentan una precisión de predicción bastante similar y elevada: 88,98% de media para el Random Forest y 88,46 % de media para el Prophet.

Respecto a la optimización de la gestión del personal se ha estimado que a pesar de que el caso que aporta más beneficios es aquel en el que se adapta el número de empleados al calculado por la herramienta, el modelo de planificación que mejor se adaptaría al negocio sería aquel en

el que se unifica el número de empleados a lo largo del turno, siendo éste el mayor valor obtenido en la optimización realizada por la herramienta.

4. Conclusiones

Este proyecto tiene un gran potencial y puede crear un gran impacto en el sector. Se demuestra que todos los escenarios de mejora propuestos incrementan sustancialmente los beneficios obtenidos para cada restaurante. Aún eligiendo el escenario menos óptimo de los propuestos, los beneficios obtenidos con el modelo de planificación actual se incrementan en un 36% para el Random Forest y en un 33,45 % para el Prophet.

Se trata de una herramienta muy completa que incluye conocimientos multidisciplinares y cuya aplicación podrá extenderse a otros sectores. Además, esta herramienta favorecerá al proceso de digitalización en el que está introduciéndose este sector, y facilitará el uso y la aplicación de algoritmos de aprendizaje supervisado e inteligencia artificial al proceso de gestión de empleados. Por otro lado, es una herramienta flexible, completa y fácil de adaptar a los clientes, pudiendo incorporar modificaciones y aspectos personalizados en cada caso.

Con esta herramienta no solo mejorarán los resultados de la empresa, también mejorará la calidad del servicio ofrecido al tener la posibilidad de personalizar la carga de trabajo de los empleados, lo cuál aumentará el nivel de satisfacción de los clientes y como consecuencia mejorará la imagen de marca de la cadena de restauración incrementando la fidelidad de los clientes. Se podrá también obtener información real sobre el nivel de ocupación de cada empleado pudiendo optimizar su rendimiento y gestionar las tareas de manera más eficiente. Además, mejorará la calidad de trabajo de los empleados y se aumentará la lealtad a la empresa y sentimiento de pertenencia, reteniendo mejor el talento.

No obstante, se han detectado una serie de mejoras del proyecto que podrían realizarse para mejorar la herramienta de cara a venderla a un cliente real o a introducirla en el mercado. Algunas de estas propuestas son: la personalización de los empleados e inclusión de sus preferencias, el refinamiento de los modelos predictivos, o el incremento de funcionalidades.

Por último, esta herramienta podrá aplicarse para otros fines como por ejemplo para proyectos de consultoría con el objetivo de simular situaciones y ofrecer propuestas de mejoras, para emprendedores que desconozcan el sector de restauración y deseen tener más información de su funcionamiento o para otros sectores como el comercio, los hoteles, centros hospitalarios... que deseen optimizar la planificación del personal.

5. Referencias

[KPMG18] KPMG. (2018). 4 | *Anuario de la restauración en España*. Retrieved from <https://assets.kpmg/content/dam/kpmg/es/pdf/2018/09/anuario-restauracion-organizada-espana.pdf>

[COSS18] Cossío, A. (2018). Bots, Machine Learning, Servicios Cognitivos Realidad y perspectivas de la Inteligencia Artificial en España, 2018. *Pwc*, 1–34. Retrieved from <https://www.pwc.es/es/publicaciones/tecnologia/assets/pwc-ia-en-espana-2018.pdf>

[CANA16] CANADEAN, “Spain Foodservice : The Future of Foodservice in Spain to 2020,” 2016.

TOOL FOR FLOW PREDICTION AND OPTIMIZATION OF EMPLOYEES PLANNING IN THE RESTORATION SECTOR.

Author: González Molina, Alejandra.

Mentor: Cruz Sánchez de Rojas, Álvaro De La.

Colaborative Entity: PricewaterhouseCoopers Asesores de Negocio S.L.

ABSTRACT

1. Introduction

The foodservice sector is currently one of the most important sectors of the Spanish economy. In addition, it is a key element in the tourism sector since it generates a great and important attraction for it.

After overcoming a long period of recession and despite having been one of the most punished sectors during that period, in 2017 the market value registered a positive evolution for the fourth consecutive year. It was in 2014 when after having bottomed out due to six consecutive years of decline in a row, a change of trend in the market was spotted.

It should be noted that the strong recovery that is taking place in the sector nowadays is mainly due to an increase in the flow traffic or in the number of visits made to the hostelry locals, and not due to an increase in the average ticket consumed [KPMG18]. Despite this, the hospitality sector as a whole has lost 14% of turnover since the crisis erupted in 2008.

On the other hand, consumer habits, tastes and behaviors, as well as the types of gastronomic experiences are in full evolution. That is the reason why companies are being forced to adapt themselves to this change by transforming their traditional business model, incorporating new technologies into it. In addition, the offers are more and more oriented to satisfying customers, who are increasingly more informed, more demanding, seek more innovation and are less loyal to brands. Furthermore, the rise of the collaborative economy and the strong penetration of smartphones are pushing this sector towards digitalization.

The key factors in this transformation that the restoration sector is experiencing are being the use of tools for data analysis, direct communication with consumers through social networks and other platforms and the development of new services that provide real added value to customers.

It should be highlighted that planning is the key to success in restaurants. Personnel management and stock management are two of the most important aspects to take into account when achieving good results. Staff is one of the most important assets of a restaurant since the number of sales will largely depend on their productivity and the service they offer. In addition, the speed of service and the quality of it will directly be translated into customer satisfaction and increasing sales.

Optimally managing the shifts and schedules of a restaurant, as well as the number of employees needed at each moment of the day, is a task of great complexity that requires a great investment of time and effort. Nowadays, restaurants are not giving to this activity the real importance it has because the impact of this activity on the results and profits of a restaurant is largely unknown. In addition, the benefits that could be obtained by an optimal management of this task are also unknown.

Currently, in the vast majority of restaurants, the most commonly used planning tool for managing the employees is Excel. Usually, general shift templates are used, where the staff needed and their corresponding schedules are assigned manually. In general, these planning methods do not take into account the possible fluctuations in restaurants fluency that may occur in the short term as a result of factors such as weather, specific events, social factors, economic factors ... to provide the exact number of personnel needed. This leads to threatening situations and serious dangers to restaurants as for instance the phenomenon known as Under / Over-Staffing problem, which refers to the excess of personnel or the lack of it.

According to some sources, the excess of hours planned in restaurants can exceed up to 15%. On the other hand, the lack of personnel at certain times can cause losses of approximately 4%. Either of these two situations involves a great loss of profits for the company, as well as a decrease in customer satisfaction and a deterioration of the brand image and reputation.

Despite the existence of certain tools for organizing staff and creating shifts, this type of tools do not allow the adaptation of the schedules and shifts of employees to the specific and foreseen needs of the restaurant, and what is more important, they usually don't incorporate external variables to the business, that have great impact on the affluence. That is to say, it does not make an analytical forecast of the needs that the establishment itself will have taking into account various factors to update the shifts and the number of employees needed in real time. On the other hand, two new tools in the market claim to use predictive analytics to estimate future affluence and plan accordingly. However, there is a great difficulty in standardizing the casuistry on a market tool, which implies that those tools can lose precision and customization possibilities face to the client, besides incurring the high costs.

2. Methodology

The main objective of the project is the development of a tool that will allow to incorporate the database of a real customer and that will be capable of using these historical data to make a prediction of affluence taking into account various factors and to calculate the optimal number of employees for each shift. The tool will incorporate supervised learning algorithms (Machine Learning) and artificial intelligence to obtain the results [COSS18].

The project will be structured in three parts: the simulation of the restoration business, the prediction of affluence for a chain of simulated restaurants and the optimization of personnel management taking into account the expected inflow for the chain.

To carry out the business simulation, the Delphi method is used. This method consists on bringing together a group of experts with whom the results obtained in the process of building the business simulation will be discussed and validated in an iterative manner until a simulation that reflects reality in the most precise way is obtained.

Once the business simulation is achieved, it will be used as a database of the simulated restoration chain. The next step will be to make the prediction of influx. To do this, two mathematical prediction models will be implemented: Random Forest and Prophet. The objective is to compare the results and verify which model offers a more accurate solution for the prediction of flow. These mathematical models will be developed, parameterized and executed in the Python programming environment.

Finally, once the estimated data for the prediction of influx of the restoration chain has been obtained, the planning of the company's employees will be optimized. For this purpose, a series

of hypotheses will be made and a series of quality standards will be established to maximize the profits of the restauration chain while offering the best possible service.

In order to demonstrate the effectiveness of this tool and the impact it can have on the market, a business case will be made and different scenarios will be analyzed offering the conclusions in this regard.

3. Results

The following figure shows the interface of the created visualization tool and presents the results obtained in each of the three mentioned parts. The first interface shows the historic of the restaurants along with the key performance indicators; the second interface includes the results of both prediction models in order to compare and verify the effectiveness and precision of each one; the third screen reflects the optimal number of employees that would be needed in each shift during the shift to maximize benefits along with the comparison of the current planning model. You can also see in this last screen a comparison of the current benefits compared to what would be obtained in the optimal case.

The three screens of the tool offer the possibility of selecting a series of filters with which the information displayed will be updated, with the aim of being able to customize the information that you want to visualize and analyze certain results in particular.

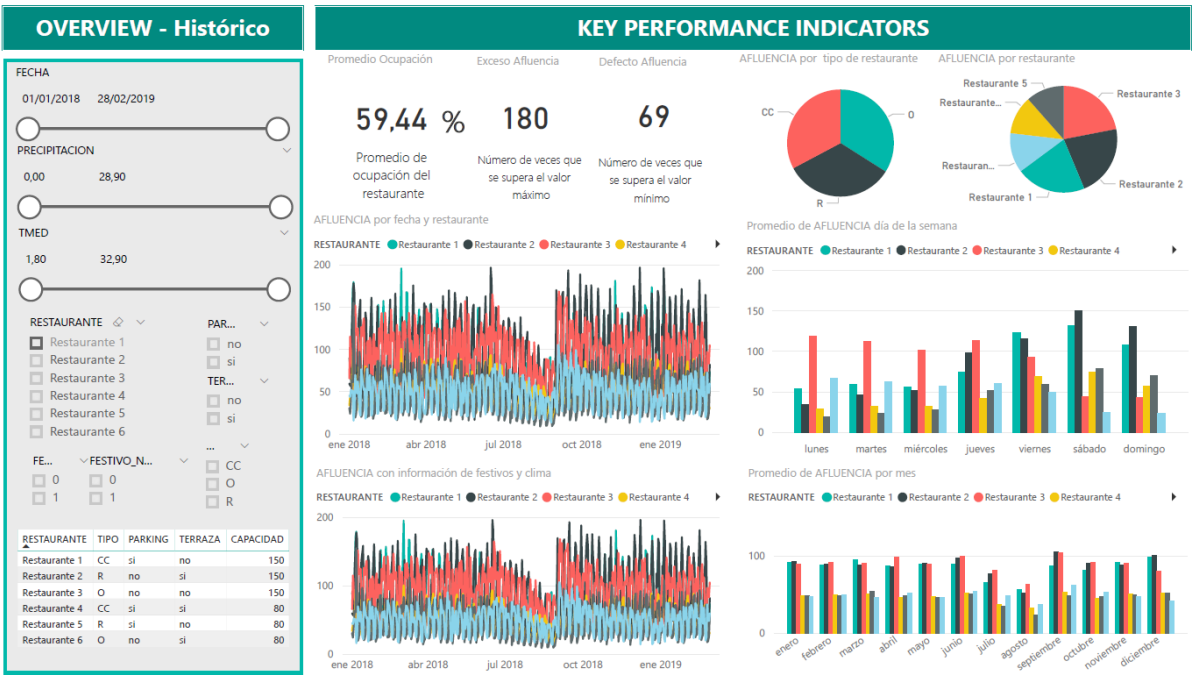


Figure 4. First interface of the tool

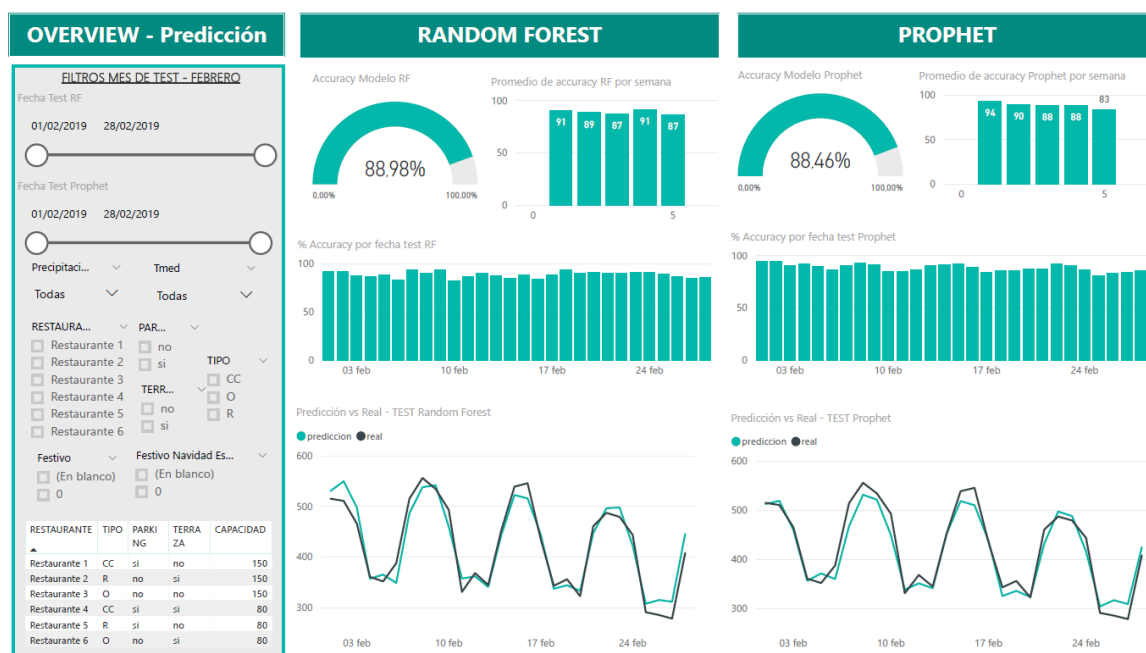


Figura 5. Second interface of the tool

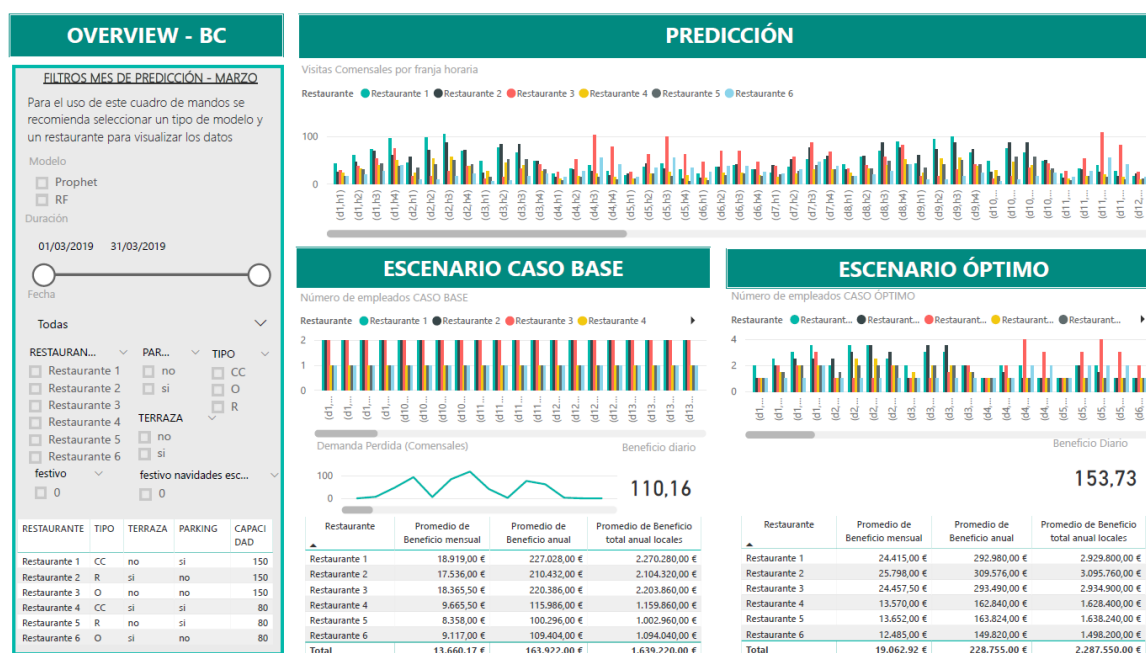


Figura 6. Third interface of the tool

The business simulation was accepted as valid after numerous iterations of correction and validation with the group of experts, including diverse factors that have direct influence on the customers flow in restaurants.

Regarding the results obtained in the prediction section, it has been observed that both models would be valid for the prediction of influx in this case since they present a fairly similar and high accuracy: 88.98% on average for the Random Forest and 88, 46% on average for the Prophet.

Regarding the optimization of personnel management, it has been estimated that although the case that provides the highest profits is the one in which the number of employees is adapted to the one calculated by the tool, the planning model that would best suit the business would be

the one in which the number of employees is unified throughout the shift. The number of employees established will be the highest value obtained in the optimization performed by the tool.

4. Conclusions

This project has great potential and can create a great impact in the sector. In addition, all the proposed improvement scenarios do substantially increase the profits obtained for each restaurant. Even choosing the least optimal scenario proposed, the benefits obtained with the current planning model are increased by a 36% for the Random Forest and by a 33.45% for the Prophet.

It is a very complete tool that includes multidisciplinary knowledge and whose application may be extended to other sectors. Furthermore, this tool will favor the digitization process in which this sector is being introduced, and will facilitate the use and application of supervised learning algorithms and artificial intelligence to the employee management process. On the other hand, it is a flexible, complete and easy to adapt to the clients tool, being able to incorporate modifications and personalized features in each case.

With this tool, not only will the company's results improve, but it will also improve the quality of the service offered by having the ability to customize the workload of employees, which will increase the level of customer satisfaction and as a consequence will improve the image of brand of the restaurant chain increasing customer loyalty. You can also get real information on the level of occupation of each employee allowing you to optimize their performance and manage tasks more efficiently. What's more, it will improve the quality of work of the employees and will increase their loyalty to the company and feeling of belonging, retaining the talent in a better way.

However, a series of propositions that could be made to improve the tool in order to sell it to a real customer or to introduce it to the market have been detected. Some of these proposals are: the personalization of employees and inclusion of their preferences, the refinement of predictive models or the increase of functionalities.

Finally, this tool can be applied to other purposes such as for consulting projects with the aim of simulating situations and offering improvement proposals, for entrepreneurs who do not know about the restoration sector and wish to have more information about its operation or for other sectors such as commerce, hotels, hospital centers ... wishing to optimize personnel planning.

5. References

- [KPMG18] KPMG. (2018). 4 | *Anuario de la restauración en España*. Retrieved from <https://assets.kpmg/content/dam/kpmg/es/pdf/2018/09/anuario-restauracion-organizada-espana.pdf>
- [COSS18] Cossío, A. (2018). Bots, Machine Learning, Servicios Cognitivos Realidad y perspectivas de la Inteligencia Artificial en España, 2018. *Pwc*, 1–34. Retrieved from <https://www.pwc.es/es/publicaciones/tecnologia/assets/pwc-ia-en-espana-2018.pdf>
- [CANA16] CANADEAN, “Spain Foodservice : The Future of Foodservice in Spain to 2020,” 2016.

Agradecimientos

Me gustaría agradecer en primer lugar a mi director de proyecto, Álvaro de la Cruz Sánchez de Rojas por su gran ayuda, disponibilidad y cercanía a lo largo de la realización del proyecto y por haber confiado en mí para realizarlo. En segundo lugar, me gustaría agradecer a Miguel García Ponce, Eduardo Ramos Rodríguez, Jacobo Cardama Aldecoa y al resto de mis compañeros de Price Waterhouse Coopers por haberme guiado, enseñado y ayudado tanto durante los meses de prácticas. Gracias por haberme dedicado parte de vuestro tiempo a pesar de vuestros cargados horarios. Trabajar con ellos me ha enseñado y aportado muchas cosas, entre ellas el valor del compañerismo y el trabajo en equipo y la importancia de gestión del tiempo. Gracias también a la Universidad Pontificia de Comillas, ICAI, por haberme dado la oportunidad de realizar prácticas en Price Waterhouse Coopers para realizar este proyecto, y en concreto al departamento de Organización por haberme guiado durante el proceso. Por último, quiero agradecer a mi familia y a mis amigos su gran apoyo a lo largo de estos años de carrera y master ya que sin ellos no habría sido posible.

Índice de contenido

1. Introducción.....	3
1.1 Sector restauración	3
1.2 Contexto del sector restauración	4
1.3 Cambios en el sector	6
1.4 Transformación Digital	7
1.5 Internet of Things (IoT), Inteligencia Artificial (IA) & Machine Learning (ML).....	8
2. Estado del arte	13
2.1 Metodología de organización de personal en restaurantes.....	13
2.2 Soluciones existentes	13
2.3 Futuras tendencias del sector (ML y IA)	14
2.4 Entorno de programación: Python	15
3. Contexto del proyecto	19
3.1 Origen del proyecto.....	19
3.2 Motivación	19
3.3 Objetivos del proyecto	20
3.4 Metodología y recursos a emplear	21
3.5 Planificación.....	22
4. Solución desarrollada.....	27
4.1 Desarrollo del modelo de simulación de afluencia	27
4.1.1 Introducción.....	27
4.1.2 Reglas de generación.....	29
4.1.3 Resultados obtenidos	32
4.1.4 Aplicación real	35
4.2 Desarrollo del modelo de predicción optimizado	37
4.2.1 Introducción.....	37
4.2.2 Modelos de predicción	38
4.2.2.1 Modelo Random Forest.....	38
4.2.2.2 Modelo Prophet.....	42
4.2.3 Análisis de los resultados de los modelos de predicción.....	43
4.2.4 Conclusiones	48
4.3 Optimización del número de empleados	49
4.3.1 Introducción.....	49
4.3.2 Desarrollo del modelo de optimización.....	49
4.3.3 Análisis de los resultados	54
4.4 Herramienta.....	59

5. Caso de negocio	67
5.1 Introducción	67
5.2 Escenarios e hipótesis	67
5.3 Estudio económico	68
5.4 Impacto y análisis de los beneficios.....	71
5.4.1 Nivel económico.....	71
5.4.2 Nivel organizacional.....	73
6. Resultados y conclusiones.....	77
6.1 Conclusiones del proyecto	77
6.2 Próximos pasos y mejoras.....	78
6.2.1 Simulación.....	78
6.2.2 Predicción.....	78
6.2.3 Optimización	79
6.3 Posibles aplicaciones.....	79
7. Bibliografía	83
8. ANEXOS	91
ANEXO A – Código Python Random Forest	91
ANEXO B – Código Python Prophet	95
ANEXO C – Importancia de variables según Random Forest.....	99
ANEXO D – Gráficas obtenidas del Modelo Prophet	105
ANEXO E – Gráficas comparativas del conjunto de testeo de ambos modelos.....	117
ANEXO F – Gráficas de distribución de los factores de afluencia.....	121

Índice de Figuras

Figura 1. Ventas Totales Restauración España	4
Figura 2. Consumos per cápita	5
Figura 3. Índice de confianza del consumidor	6
Figura 4. Procesos en los cuáles se aplica el análisis de datos.....	7
Figura 5. Capacidades de los sistemas de Inteligencia Artificial (IA).....	9
Figura 6. Incremento del PIB mundial en 2030 (trillones \$)	9
Figura 7. Proceso de elaboración de un modelo de predicción.....	22
Figura 8. Simulación afluencia Restaurante 1	32
Figura 9. Simulación afluencia Restaurante 2.....	32
Figura 10. Simulación afluencia Restaurante 3.....	33
Figura 11. Simulación afluencia Restaurante 4.....	33
Figura 12. Simulación afluencia Restaurante 5.....	33
Figura 13. Simulación afluencia Restaurante 6.....	33
Figura 14. Promedio de afluencia semanal	34
Figura 15. Promedio de afluencia mensual	34
Figura 16. Promedio de afluencia en restaurantes de tipo “Comida rápida”	34
Figura 17. Promedio de afluencia en restaurantes de tipo “Comida tradicional”	35
Figura 18. Árboles de decisión en Random Forest	39
Figura 19. Random Forest Simplificado	40
Figura 20. Test Random Forest Restaurante 1	45
Figura 21. Error de predicción en el conjunto de datos de testeo Restaurante 1 (RF))	45
Figura 22. Test Prophet Restaurante 1	46
Figura 23. Predicción afluencia Restaurante 1 (Prophet).....	46
Figura 24. Propiedades de los datos (Prophet).....	47
Figura 25. Factores de afluencia para el Restaurante 1 (localizado en centro comercial).....	49
Figura 26. Factores de afluencia para el Restaurante 2 (localizado en barrio residencial).....	49
Figura 27. Factores de afluencia para el Restaurante 3 (localizado en zona de oficina)	50
Figura 28. Factores de afluencia para el Restaurante 4 (localizado en centro comercial).....	50
Figura 29. Factores de afluencia para el Restaurante 5 (localizado en barrio residencial).....	50
Figura 30. Factores de afluencia para el Restaurante 6 (localizado en zona de oficinas).....	50
Figura 31. Visualización gráfica de factores de afluencia para el Restaurante 1.....	51
Figura 32. Resultado optimización Restaurante 1.....	55
Figura 33. Resultados optimización Restaurante 6	56
Figura 34. Primera interfaz de la herramienta (Histórico)	61
Figura 35. Segunda interfaz de la herramienta (Predicción).....	62
Figura 36. Tercera interfaz de la herramienta (Optimización empleados).....	63
Figura 37. Tercera interfaz de la herramienta aplicando filtros	64

Índice de Tablas

Tabla 1. Caracterización de restaurantes.....	28
Tabla 2. Caracterización de calendario simulación.....	28
Tabla 3. Factores regla diaria	29
Tabla 4. Factores regla mensual.....	30
Tabla 5. Factores regla eventos y festivos	30
Tabla 6. Factores regla climatológica	31
Tabla 7. Factores de regla estacionalidad	31
Tabla 8. Factores de regla vacaciones escolares	32
Tabla 9. Resumen de precisión de los modelos para cada restaurante.....	44
Tabla 10. Hipótesis de costes asumidos para calcular beneficios	67
Tabla 11. Beneficios caso base con ratio de 15 comensales por empleado	68
Tabla 12. Beneficios caso base con ratio de 20 comensales por empleado	69
Tabla 13. Beneficios caso óptimo con ratio 15 comensales por empleado.....	69
Tabla 14. Beneficios caso óptimo redondeando el número de empleados	70
Tabla 15. Beneficios caso óptimo aumentando el ratio de comensales por empleado a 20.	71
Tabla 16. Resumen beneficios mensuales de cada escenario para datos del Random Forest....	72
Tabla 17. Resumen beneficios mensuales de cada escenario para datos del Random Forest....	73

CAPÍTULO I : INTRODUCCIÓN



1. Introducción

En primer lugar, es preciso definir el ámbito de estudio de este proyecto. El objetivo principal del proyecto es la elaboración de una herramienta de trabajo para restaurantes que incluya un modelo de predicción de afluencia que permita optimizar la gestión de recursos y en concreto, la gestión de personal. Para comprender el desarrollo del proyecto es preciso analizar el sector en el que se llevará a cabo y el contexto en el que surge el mismo.

Resulta importante distinguir entre sector de hostelería y sector de restauración. La hostelería es el sector formado por dos ramas de actividad, restauración y alojamiento. En este caso, el proyecto se centrará en el sector de restauración.

1.1 Sector restauración

El sector restauración se trata de uno de los sectores más importantes para la economía española. Además, es un elemento clave en el sector turístico ya que genera un gran atractivo para el mismo.

La *Federación Española de Hostelería (FEHR)* [FIGU17] realiza una clasificación del sector según los tipos de establecimientos que lo componen. Estos establecimientos pueden ser: *restaurantes*, que a su vez se clasifican en gama alta, precio medio, económico, o restaurantes rápidos de take away; *bares y cafeterías*, repartidas en bares restaurantes, bares y cervecerías, cafeterías y chocolaterías, horchaterías...; o *restauración social* formada por enseñanza, militar, residencias, sanidad y deportes.

Al mismo tiempo, se pueden identificar dos segmentos de actividad que nos permitirán comprender con mayor claridad el sector. Los segmentos son los siguientes:

- Restaurantes con servicio en mesa. Clasificados a su vez en *cafeterías*, *cervecerías* y/o *tapas*, *cocina española* y *cocina extranjera*.
A su vez, se podrían identificar los siguientes sub segmentos de actividad según el modo de gestión establecido y la oferta de restauración.
 - Restauración tradicional
 - Restauración informal. Formado por cadenas de restaurantes con aspecto moderno y original, ambiente informal y precio ajustado.
- Restaurantes sin servicio en mesa. Clasificados en: *Autoservicios tradicionales*, excluyendo aquellos establecimientos de comida rápida que tienen formato de autoservicio y *comida rápida*.

Por otro lado, la conocida como *Marcas de Restauración* se trata de una asociación empresarial formada por las 40 principales cadenas de restauración organizada y moderna, incluyendo casi 100 marcas. La restauración organizada está formada por grandes cadenas de establecimientos bien en propiedad o en franquicia, que cuentan con un modelo de gestión centralizada y avanzada. Además, cuentan con una gran capacidad de captar las nuevas tendencias de los consumidores. Servirá de referencia para el estudio del sector.

1.2 Contexto del sector restauración

Tras superar un largo período de crisis y a pesar de haber sido uno de los sectores más castigados, en 2017 el valor del mercado registró una evolución positiva por cuarto año consecutivo. Fue en 2014 cuando tras haber tocado fondo después de seis ejercicios consecutivos a la baja, se produjo un cambio de tendencia en el mercado. Esta evolución se ve contextualizada en un aumento del dinamismo del consumo en hogares a la vez que un fuerte crecimiento del turismo extranjero. Todo ello ha provocado un impulso en las ventas.

Cabe destacar que la fuerte recuperación que se está produciendo en el sector en estos momentos es debida a un aumento en el tráfico o número de visitas realizadas a los locales de hostelería, y no a un aumento en el ticket medio [KPMG18]. A pesar de ello, el sector de la hostelería en conjunto ha perdido un 14% de facturación desde que estalló la crisis en 2008.

La facturación del sector de restaurantes en el año 2017, según el informe elaborado por el Observatorio Sectorial DBK de INFORMA [DBKI17], ha sido cercana a los 23.100 millones de euros, lo que supone un 4,7% más que en 2016, que al mismo tiempo aumentó un 5,0% respecto a 2015. Sin embargo, según el Anuario de la restauración organizada en España [KPMG18] elaborado por KPMG, la facturación agregada de las empresas asociadas a Marcas en ese mismo año a pesar de haber crecido un 6,5%, experimentó un resultado ligeramente inferior al del 2016. Se estima que ello puede estar asociado a la bajada de facturación de Cataluña, que ha tenido un gran impacto en el sector, como consecuencia de la inestabilidad política y de la incertidumbre.

Se observa en la Figura 1 que la evolución de ventas ha sido también positiva en los últimos años. En 2016 se registraron 35.146 millones de euros en ventas, lo cual supuso un incremento del 1,9% respecto al año anterior. En 2017 las ventas del sector aumentaron un 2,5% alcanzando así 36.024 millones de euros, según el informe de The NPD Group. Sin embargo, los valores de ventas actuales se encuentran todavía por debajo de los 42.413 millones alcanzados en 2008.

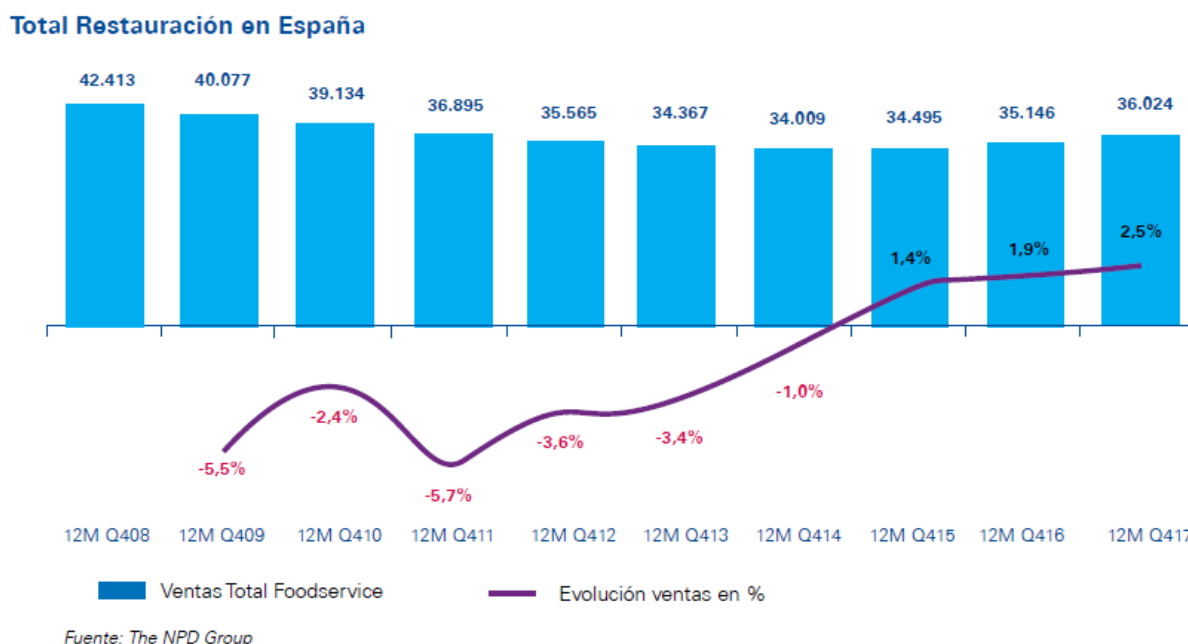


Figura 1. Ventas Totales Restauración España

Tal y como se ha mencionado previamente, este incremento se debe en gran parte al incremento del 1,3% en el número de visitas, alcanzando un valor de 98 millones de consumos adicionales

fuera de casa. Paralelamente, se ha detectado una subida del 1% en el gasto medio por cada comensal, obteniendo una media de 4,88 euros por visita realizada. Dato muy similar al registrado en 2008, donde el gasto medio era de 4,93 euros. Podría decirse que el gasto medio está estancado. Respecto al año anterior la subida fue de un 0,4 % hasta alcanzar los 4,82 euros.

Se puede observar en la Figura 2 que el ritmo de crecimiento durante los últimos tres años es constante. Aun así, quedarían 27 consumos anuales para alcanzar los niveles establecidos en 2008, donde se alcanzó un máximo histórico.

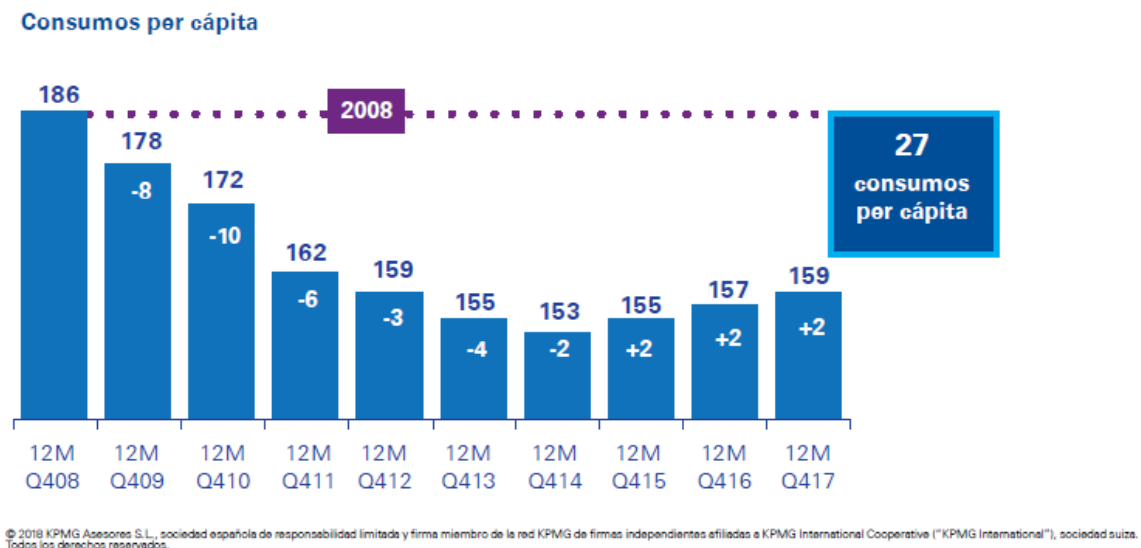


Figura 2. Consumos per cápita

Cabe destacar que el nuevo motor principal de tráfico en el sector, aportando un 86% del aumento de visitas registrado, son los restaurantes de servicio rápido. Además, según Vicente Montesinos, director ejecutivo de NPD España, es el canal organizado el que está impulsando el crecimiento en el sector.

Por otro lado, en 2017 los momentos de consumo que han generado mayores ganancias de tráfico de clientes han sido los desayunos y almuerzo. Ambos se han concentrado entre semana. Este cambio está relacionado con la recuperación del mercado laboral. Durante el fin de semana el tráfico ha aumentado a la hora de la comida sobretodo en restaurantes a la carta.

Existe una fuerte correlación entre el PIB y el consumo privado con el sector de restauración. En concreto, la restauración en España representa el 6,2% del PIB nominal. Por otro lado, la restauración organizada comercial representa el 0,51% del PIB. Para el año 2018, las previsiones de crecimiento según la Federación Española de Hostelería (FEHR) se encuentran alrededor del 4,8%. Las previsiones para las Marcas de Restauración oscilan en torno a 6,1%. En 2017 se pudo observar claramente el impacto directo que tiene el crecimiento del país en el mercado de restauración. En 2017 se registró un incremento del 3% del PIB a la vez que un aumento de 400.000 nuevos empleos al año, lo que aceleró el ritmo de crecimiento del año anterior incrementando las ventas en un 2,5% y las visitas en un 1,3%.

Otro factor que tiene influencia en el sector es el del índice de confianza del consumidor, que muestra el cambio de tendencia de perspectiva de los consumidores. Tal y como se observa en la siguiente figura, el índice de confianza del consumidor en el mes de septiembre ha descendido 11,9 puntos por debajo del mes anterior. Esto puede ser debido a la caída de 16,3

puntos de la valoración de la situación actual. Este cambio se verá reflejado probablemente en el consumo registrado en dicho mes.

	Sep	Oct	Nov	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep
	2017	2017	2017	2017	2018	2018	2018	2018	2018	2018	2018	2018	2018
ÍNDICE CONFIANZA CONSUMIDOR	103,2	99,6	100,5	102,5	102,3	99,7	98,6	99,9	97,7	107	106,1	102,4	90,6
ÍNDICE DE SITUACIÓN ACTUAL	95,8	93,4	91,2	92,9	93,5	90,6	90,1	92,7	91,3	102	101,7	99,3	82,9
ÍNDICE DE EXPECTATIVAS	110,5	105,8	109,7	112,0	111,1	108,7	107,2	107,2	104,1	112	110,4	105,6	98,2

Figura 3. Índice de confianza del consumidor

A modo de resumen, el mercado español generó 880 millones de euros adicionales en 2017 tras haber recuperado aproximadamente unos 100 millones de visitas y tras haber aumentado ligeramente el gasto medio. Aun así, en estos últimos tres años tan solo se ha recuperado una quinta parte del mercado que se alcanzó en 2007. Esto supone una gran oportunidad de recuperación, estimada en más de 6.000 millones de euros adicionales.

1.3 Cambios en el sector

Los hábitos de consumo, los gustos y los comportamientos de los clientes, así como los tipos de experiencias gastronómicas se encuentran en plena evolución. Es por ello que las empresas se están viendo obligadas a adaptarse a este cambio transformando su tradicional modelo de negocio, incorporando las nuevas tecnologías en él.

Las ofertas cada día están más orientadas a satisfacer a los clientes, que cada vez están más informados, son más exigentes, buscan más innovación y son menos fieles a las marcas. A su vez, el auge de la economía colaborativa y la fuerte penetración de los *smartphones* están empujando a este sector hacia la digitalización.

Se pueden apreciar cambios en el sector en diversos aspectos. Por ejemplo, se está produciendo un cambio en los tipos de consumición y de servicio. Este cambio se puede apreciar en la evolución de menús de los restaurantes. Los clientes buscan innovación y cada vez están más orientados hacia un consumo más saludable ya que ha aumentado la preocupación por el factor salud. Además, cada vez más se busca la personalización individual de los menús en tiempo real.

Por otro lado, la aparición del nuevo tipo de consumidores conocido como *millennials*, está acelerando los procesos de transformación debido al factor tecnológico que arrastran con ellos y al tipo de consumo que llevan a cabo.

El tipo de consumo también ha evolucionado ya que el tiempo dedicado a las comidas se ha visto reducido drásticamente en la última década. Lo cual ha impulsado una nueva rama de restauración conocida como el “*grab and go*” y servicios de distribución de comida preparada.

Los clientes de hoy en día buscan experiencias digitales para conectar con los restaurantes. El uso de las redes sociales ha impulsado la creación de programas de fidelización y de marketing digital para establecer conexiones entre clientes y restaurantes.

Por último, ha aumentado la preocupación por la sostenibilidad y el medio ambiente. Este factor afecta también al sector restauración ya que gran parte de los restaurantes han tenido que adaptar sus negocios a estas preocupaciones para poder seguir en el mercado. Las cadenas de producción y distribución se han visto también afectadas por estos cambios.

1.4 Transformación Digital

El sector de restauración se enfrenta a un nuevo entorno de relación con los clientes que está marcado por la digitalización y por el impulso de los avances tecnológicos y herramientas que amplifican la interacción entre usuarios y empresas.

Las empresas disponen cada vez de más información de los clientes y pueden utilizarla en su beneficio para conocer qué están esperando exactamente para poder ajustar la oferta a sus deseos. De este modo pueden anticiparse a posibles cambios en los hábitos de consumo y personalizar tanto la oferta como el producto.

La forma de comunicación entre los clientes y las empresas también está evolucionando gracias a las redes sociales e internet en general. Los propios consumidores tienen un enorme peso en las decisiones de las empresas, así como gran impacto en los resultados de las mismas.

Por otro lado, este fenómeno de digitalización también está provocando la aparición de nuevos servicios en la cadena de valor de la restauración. Estos servicios están directamente ligados a las nuevas tecnologías. Estas tecnologías pueden tener distintos objetivos. Por un lado, pueden estar ligadas a la automatización o transformación de ciertas actividades para la simplificación de tareas o mejora de servicios; y por otro lado al análisis y previsión de datos.

En conclusión, los factores claves en esta transformación del sector de restauración están siendo el uso de herramientas para el análisis de datos, la comunicación directa con los consumidores a través de redes sociales y otras plataformas, y el desarrollo de nuevos servicios que aportan verdadero valor añadido a los clientes. Según el *Anuario de la restauración organizada en España* (KPMG, 2018) elaborado por KPMG los procesos o funciones para los cuáles las empresas están utilizando el análisis de datos son los siguientes [KPMG18]:

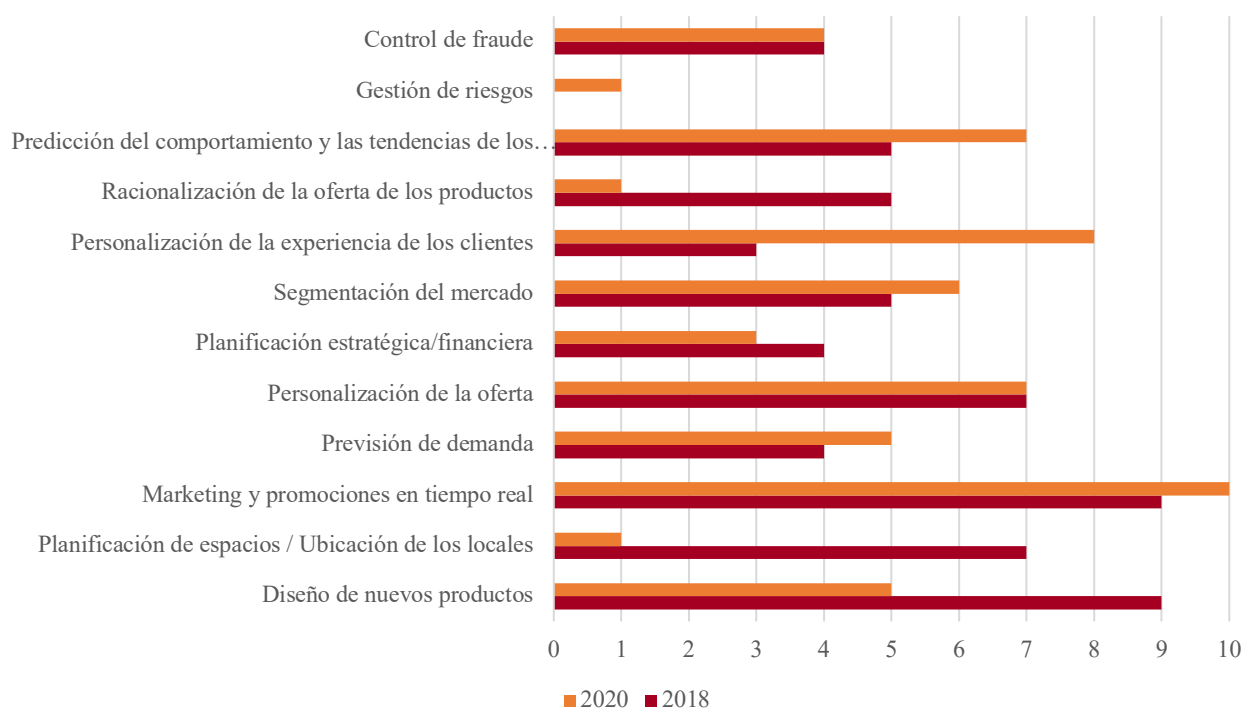


Figura 4. Procesos en los cuáles se aplica el análisis de datos

Además de los factores mencionados previamente, cabe destacar que a día de hoy la transformación se manifiesta principalmente en los consumidores. El 46% de los españoles

eligen el restaurante a través de internet y más del 70% inicia su relación con los mismos en entornos digitales, conocidos como “umbrales digitales”. El 43% de los consumidores utilizan el teléfono móvil durante la comida. El 60% de los clientes dejan una evaluación o comentario sobre el restaurante en las redes sociales. Sin embargo, a pesar de ser uno de los principales sectores afectados por esta revolución, se estima que tan solo un 65,6% de los restaurantes disponen de página web propia.

Según Marius Robles, CEO de Reimagine Food, en el futuro existirán tres grandes tipos de restaurantes que cubrirán distintas necesidades de consumo:

- *Restaurantes inteligentes o de conocimiento*, donde será posible experimentar nuevas emociones desde una perspectiva multidimensional. Estos restaurantes contarán con un alto componente tecnológico, con cocinas conectadas con los proveedores habituales.
- *Restaurantes de rendimiento*, de tipo funcional y centrados en la personalización. En ellos reinará el concepto “healthy” y “wellness”.
- *Restaurantes de placer*, donde el único objetivo será el de disfrutar.

1.5 Internet of Things (IoT), Inteligencia Artificial (IA) & Machine Learning (ML)

El Internet de las Cosas, también conocido como *Internet of Things (IoT)*, se trata de una red de objetos físicos que utilizan una serie de sensores y APIs para conectarse e intercambiar datos a través de internet. Dentro de las tecnologías que hacen posible IoT encontramos las herramientas de gestión de Big Data, la Inteligencia artificial y Machine Learning, las analíticas predictivas, la nube y la identificación por radiofrecuencia o RFID. La información en forma de Big Data que produce el IoT podrá traducirse a datos que ayudarán a predecir acontecimientos y comportamientos, así como a ejecutar tareas de manera automática. Sin embargo, este intercambio de información no aportará valor alguno a las empresas si no se utilizan otras tecnologías para extraer conclusiones de estos datos y poder utilizarlos. Para ello utilizaremos tecnologías como la Inteligencia Artificial.

Inteligencia Artificial es el concepto que incluye todas aquellas tareas que son realizadas por programas, softwares o máquinas y que harían uso de la inteligencia humana si fueran realizadas por personas. Se trata de una herramienta que permite administrar la información que ha sido generada por el IoT.

Por otro lado, Machine Learning es una aplicación actual de la Inteligencia Artificial que se basa en dar a las máquinas acceso a los datos y hacer que aprendan por sí mismos a través del uso de algoritmos de análisis de datos. A través de ellos podrá aprender de los mismos datos y podrá hacer una determinación o predicción sobre algo en el mundo. En la siguiente figura apreciamos las capacidades de los sistemas de IA, que intentan replicar el procesamiento del cerebro humano en los siguientes aspectos:

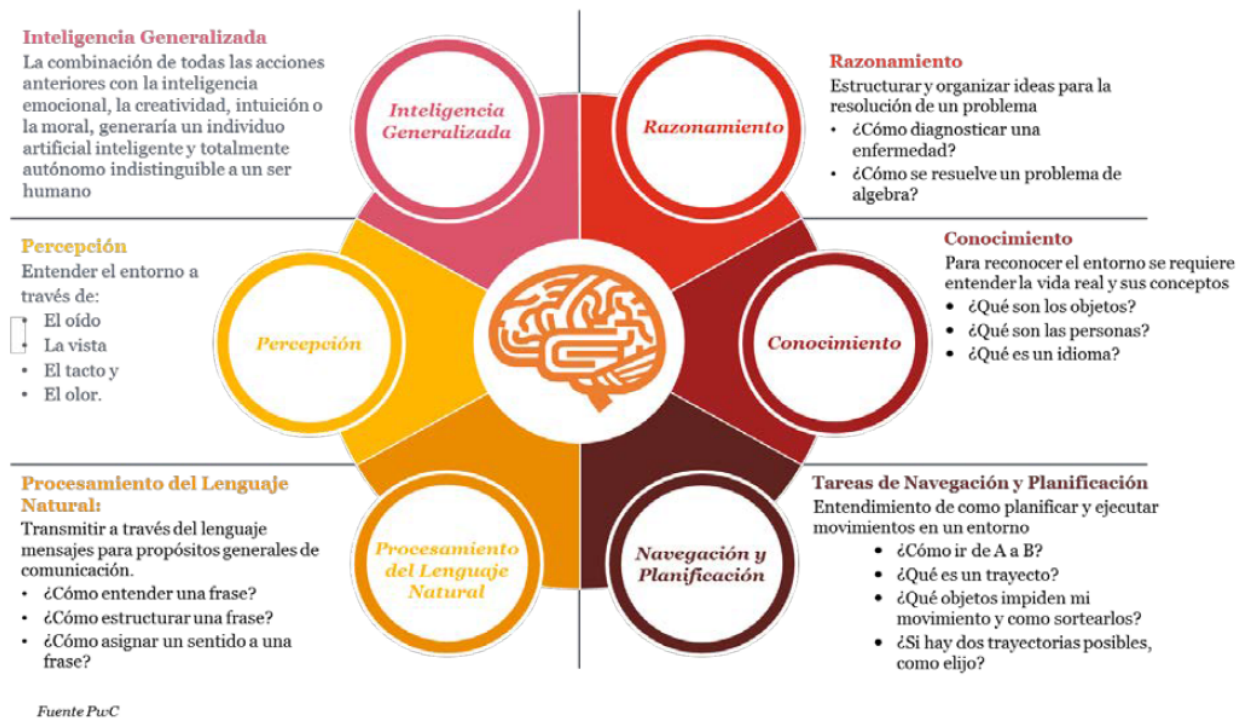


Figura 5. Capacidades de los sistemas de Inteligencia Artificial (IA)

Estos nuevos avances tecnológicos son el motor del cambio que se implementará para responder a las nuevas necesidades de los clientes. El uso de estas tecnologías nos permitirá gestionar la gran cantidad de información a la que tenemos acceso a día de hoy, permitiendo la toma de decisiones estratégicas a partir de los datos analizados mediante la Inteligencia Artificial, así como ofrecer un trato al cliente a nivel más personalizado. Cabe destacar que, entre negocios, consumidores y gobiernos, se prevé que en 2020 se invertirán alrededor de 1,6 trillones de dólares en instalar soluciones de IoT. Un estudio realizado por Price Waterhouse Coopers sobre IA e IoT afirma que, en 2030, el PIB mundial puede incrementar hasta un 14% debido a la Inteligencia Artificial [COSS18].

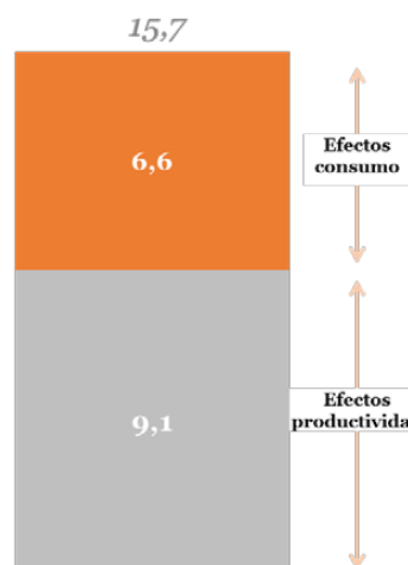


Figura 6. Incremento del PIB mundial en 2030 (trillones \$)

En cuanto a la aplicación de estas tecnologías al sector de restauración, a pesar de que se trata de uno de los sectores más expuestos a la digitalización como consecuencia del perfil de los clientes y la creciente utilización de las redes sociales, aún queda mucho por camino por recorrer. Aun así, se está acercando a las compañías que lideran el uso eficiente de datos, tanto para conocer mejor al cliente y poder ofrecer ofertas personalizadas (*front-office*), como para hacer los procesos y planificaciones más eficientes (*back-office*). Más adelante se presentarán algunas de las tecnologías utilizadas actualmente y aquellas que están empezando a aparecer.

A día de hoy, las principales herramientas digitales que ayudan a agilizar la actividad en estos negocios son los TPV o Terminal de Punto de Venta utilizados por el 85,3% de los establecimientos, los comanderos electrónicos, utilizados por un 26,1% de los restaurantes, y otro tipo de herramientas como avisadores, datafonos, Aj Cash ... utilizadas por un 2,5% del sector. No obstante, aún existe un 12,4% de empresarios que no disponen de ninguna herramienta digital.

Por otro lado, cada vez son más comunes los sistemas digitales de gestión y planificación dentro de este sector. Aproximadamente un 45% de los negocios de restauración cuenta con este tipo de herramientas. Alrededor de un 40% de restaurantes cuenta con un sistema de gestión de reservas mientras que un 55% de los restaurantes disponen de un sistema de reputación online.

Según los profesionales del sector, el restaurante del futuro será cada vez más personalizado, inteligente y digital, ya que el cliente del siglo XXI tiene como objetivo la búsqueda de experiencias gastronómicas, vanguardistas, innovadoras y personalizadas. Respecto a la transformación digital en restauración, aún hay que trabajar para impulsar soluciones óptimas que mejoren la competitividad del sector y ayuden a la captación y fidelización de clientes, controlando mejor los ingresos y los gastos.

CAPÍTULO II : ESTADO DEL ARTE



2. Estado del arte

2.1 Metodología de organización de personal en restaurantes

La planificación es la clave del éxito en los restaurantes. La gestión del personal, así como la gestión del stock son dos de los aspectos más importantes a tener en cuenta a la hora de lograr buenos resultados. El personal es uno de los activos más importantes de un restaurante ya que el número de ventas dependerá en gran parte de su productividad y del servicio que ofrezcan. La rapidez de servicio y la calidad del mismo se traducen directamente en satisfacción del cliente.

Gestionar de manera óptima los turnos y horarios de un restaurante, así como el número de empleados necesarios para cada momento del día se trata de una labor de gran complejidad que requiere una gran inversión de tiempo y esfuerzo. A día de hoy, no se le otorga a esta actividad la importancia que realmente tiene ya que en gran parte se desconoce el impacto de la misma sobre los resultados de un restaurante, así como los beneficios que se podrían obtener por una óptima gestión de esta tarea.

Actualmente, en la gran mayoría de restaurantes, cafeterías y bares la herramienta de planificación de personal más utilizada es Excel. Normalmente se utilizan plantillas de turnos generales, en las cuáles se va asignando manualmente el personal necesario y los horarios correspondientes. En general, en estos métodos de planificación no se tienen en cuenta las posibles fluctuaciones que pueden llegar a ocurrir en un corto plazo como consecuencia de factores como el tiempo atmosférico, acontecimientos puntuales, factores sociales, factores económicos... para prever el personal necesario. Lo cual puede llevar a situaciones amenazadoras y de grave peligro para los restaurantes como el fenómeno conocido como *Under/Over-Staffing*, que hace referencia al exceso de personal o la falta del mismo. Según algunas fuentes, el exceso de horas planificadas en restaurantes puede llegar a superar el 15%. Por otro lado, la falta de personal en ciertos horarios puede provocar pérdidas de aproximadamente el 4%. Cualquiera de estas dos situaciones supone una gran pérdida de beneficios para la empresa, así como una disminución en la satisfacción del cliente que puede conllevar grandes pérdidas a corto y largo plazo.

Planificar el personal de un restaurante no solo consiste en calcular los horarios de cada turno y asignar trabajadores a los mismos. También es preciso reflexionar sobre las horas picos que se producen a lo largo del día y aquellas en las que habrá menos tráfico de clientes, qué perfil de empleado es óptimo en cada caso, qué tareas se deben realizar en cada caso, qué tipo de contrato tiene cada empleado, que restricciones laborales hay que cumplir... Por otro lado, es preciso tener en cuenta y respetar los días libres y las vacaciones de cada empleado a la vez que gestionar las bajas que se producen. Al utilizar la herramienta de Excel todos estos cálculos se realizan generalmente de forma manual. Además de ser poco eficiente, no se alcanza el punto óptimo y por lo tanto hay un margen de mejora que podría traducirse en mayores beneficios.

2.2 Soluciones existentes

Actualmente existen ciertos programas o softwares que han sido creados para gestionar la planificación del personal en restaurantes. Se trata de soluciones estándar que permiten la generación automática de los turnos de empleados, eliminando así el proceso manual utilizado previamente con Excel. Este tipo de programas garantiza el cumplimiento de las regulaciones. No obstante, estas herramientas no son capaces de realizar ningún tipo de análisis de datos.

Normalmente estos softwares de planificación se encuentran integrados dentro del propio programa de gestión del restaurante, donde se puede controlar la facturación a tiempo real, las

necesidades de stock, la situación de almacenaje y reposición, el control de las mesas y terraza, la comunicación con cocina, el flujo de salida de platos ...

Algunos de las empresas que elaboran los programas de gestión de personal más utilizados en el mercado de restauración a día de hoy son [GARC17]:

- *Abantime*, consultora informática que cuenta con una amplia variedad de tipos de software según el tipo de restaurante (caterings, fast foods, cafeterías, pizzerías ...). La planificación de personal se realiza a través de su teléfono inteligente y a través de la tablet.
- *Hosteltactil*, empresa que ofrece programas de gestión de TPV. El software de *Hosteltactil* se caracteriza por la sencillez de utilización y su interfaz dinámica.
- *Cuiner*, empresa que ofrece tres tipos de software según el tipo de negocio que lo solicite (pequeños negocios, empresas de restauración consolidadas y cadenas de restaurantes). Este software ofrece como valor añadido la elaboración de informes y estadísticos sobre el local.
- *Restaweb y techni-web*, aplicación premiada por la Asociación Madrileña de Empresarios de la Restauración como el Mejor Software de Restauración. Esta aplicación ofrece distintos niveles de utilización según el rol que tenga el usuario dentro de la empresa (camareros, gerentes, cocineros...). Es posible acceder a la aplicación desde cualquier lugar, sin necesidad de estar dentro del propio establecimiento y controlar el mismo a distancia. Además, facilita la coordinación de operaciones entre sala, cocina y gestión administrativa.
- *Ofibarman*, software flexible y adaptable a las necesidades del negocio. Permite controlar el TPV, realiza la comunicación entre cocina y camareros, y controla el stock disponible. Además, es posible realizar la planificación del personal según la configuración introducida.
- *Ofimesas*, cuenta con un programa de marketing, administración y gerencia de gran potencia. Es una buena herramienta para la gestión y el control de reservas.
- *XD Software*, especializado en comida rápida, permite gestionar los pedidos y optimizar las reservas.

Sin embargo, este tipo de herramientas no permite la adaptación de los horarios y turnos de los empleados a las necesidades del restaurante. Es decir, no realiza una previsión analítica de las necesidades que va a tener el propio establecimiento teniendo en cuenta diversos factores para actualizar los turnos y número de empleados necesarios en tiempo real.

2.3 Futuras tendencias del sector (ML y IA)

Teniendo en cuenta la información presentada previamente, cabe destacar el importante papel que van a desarrollar las nuevas tecnologías y en concreto, las tecnologías ligadas a la explotación de datos, en la planificación de personal en el sector de restauración [KPMG18]. Por lo tanto, las futuras tendencias del sector se encuentran en la Inteligencia Artificial y el Machine Learning como herramientas para el análisis de datos que permitan la optimización y gestión del personal.

Las soluciones utilizadas a día de hoy son incapaces de predecir la demanda que habrá en cada restaurante para dimensionar la plantilla de forma óptima. No obstante, se están empezando a desarrollar nuevos softwares que implementan sistemas de analítica avanzada y predicción de demanda. Este tipo de programas son capaces de generar horarios inteligentes cuyo objetivo es

la búsqueda de la rentabilidad máxima del establecimiento. La satisfacción del cliente y de los empleados son objetivos primordiales para estas herramientas, ya que conducirán a mayores beneficios.

Estas nuevas herramientas son capaces de predecir la cantidad de visitantes que tendrá el establecimiento en cada momento del día, y por lo tanto estimar el número óptimo de personal que es necesario para abastecer dicha demanda. Como resultado, el personal estará perfectamente dimensionado en cada momento del día y podrá realizar sus tareas asignadas en la mejor de las situaciones. Esto permitirá mejorar la experiencia del cliente, el cual divulgará buenas críticas, que se verá traducido en atracción de nuevos comensales.

Algunos de los softwares que están empezando a utilizar la inteligencia artificial y analítica avanzada, pero cuya eficacia real está aún por demostrar son:

- *Orquest*, consiste en una solución que permite predecir la demanda, visitas y ventas que tendrá el restaurante según los distintos momentos del día, hora a hora. Una vez analizados los datos, permite la elaboración de una plantilla de trabajo que se ajusta a las necesidades del restaurante, y que tiene en cuenta las capacidades y tareas de los propios trabajadores. Además, permite la participación activa de los empleados en la gestión del horario, permitiéndoles expresar sus preferencias a través de una aplicación móvil. Por último, permite medir el rendimiento y la productividad del establecimiento para identificar los puntos débiles que se deben mejorar.
- *GIRnet (Mapal Software)*, se trata de un software de gestión e inteligencia de negocio adaptado a las necesidades del sector. Esta aplicación genera turnos óptimos, ajustando el personal a la demanda real, utilizando algoritmos de Inteligencia Artificial. A través de la propia aplicación es posible comunicar las preferencias de turnos y realizar cambios en tiempo real. También es posible controlar mediante huella dactilar los tiempos de trabajo. La herramienta es capaz de realizar diversos análisis comparativos. Esta solución permite aumentar las ventas y productividad a la vez que mejora el nivel del servicio.

El uso de este tipo de software permitirá reducir el coste laboral, incrementar la satisfacción del cliente y del empleado, aumentar la productividad y aumentar los beneficios. Puesto que la implementación y utilización de dichos programas se encuentra en las primeras etapas, su eficacia real aún está por demostrar. Además, existe una gran dificultad a la hora de estandarizar las casuísticas sobre una herramienta de mercado, por lo que es posible que dichas herramientas puedan perder precisión y capacidad de personalización de cara al cliente, además de incurrir en elevados costes. A pesar de poder predecir la afluencia para restaurantes, es difícil que estas herramientas incluyan factores externos al negocio y particulares al cliente en cuestión al tratarse de una herramienta más estandarizada para poder aplicarla a diversos clientes.

2.4 Entorno de programación: Python

La base de muchos de los softwares presentados anteriormente es el lenguaje de programación Python. Se trata de un lenguaje de propósito general, multiplataforma y multiparadigma, de programación libre y de código abierto. Por lo general se trata de un código fácilmente legible y limpio. Python soporta la programación orientada a objetos.

Dicho lenguaje puede utilizarse para distintas funciones como por ejemplo para el desarrollo web, para la creación de un servidor, para la extracción de datos, para la informática científica e incluso para el aprendizaje automático.

Una de las ventajas de Python es que permite crear tanto sitios web como aplicaciones en una amplia gama de sistemas operativos que incluyen entre otros: Windows, Mac, iOS y Android. Además, es perfecto para trabajar con volúmenes de datos de gran tamaño ya que favorece su extracción y procesamiento.

Por otro lado, al ser un lenguaje interpretado no necesita compilar el código fuente para poder ejecutarlo. Esto supone una gran ventaja ya que aumenta la rapidez de desarrollo. Además, contiene un gran número de librerías, tipos de datos y funciones que permite realizar ciertas tareas sin tener que programar desde el inicio.

Python dispone también de un intérprete por línea de comandos que permite introducir sentencias y ejecutarlas individualmente de forma que se pueda visualizar el resultado. Esto hace posible probar los resultados de la ejecución de partes de código de manera sencilla y rápida.

Frente a otros lenguajes de programación utilizados actualmente, Python destaca por su sencillez. El número de líneas de código utilizadas para crear un mismo programa suele ser menor en Python que en programas como Java o C por ejemplo.

Todo ello hace que este lenguaje sea óptimo para la creación de herramientas destinadas al análisis y predicción de datos, con el objetivo de optimizar recursos. Como es el caso del proyecto que se presenta en los siguientes apartados.

CAPÍTULO III : CONTEXTO DEL PROYECTO



3. Contexto del proyecto

3.1 Origen del proyecto

La propuesta inicial de este proyecto fue realizada por la firma de auditoría y consultoría *Price Waterhouse Coopers*. En concreto Álvaro de la Cruz Sánchez, Manager del departamento de Operaciones y Supply Chain en Consultoría de Negocio y director del proyecto, detectó la necesidad por parte de las cadenas de restauración de mejorar la gestión del personal en sus distintos locales.

El problema principal identificado fue que ante un mismo evento inesperado como por ejemplo fuertes precipitaciones o eventos sociales particulares, los locales de una misma cadena de restauración sufrían diversas consecuencias, que se derivaban en importantes pérdidas. En algunas ocasiones dichos eventos provocaban falta de personal en los locales, lo cuál afectaba de manera negativa a la calidad del servicio, aumentos de tiempo de espera y deterioro de la imagen de marca de la cadena. En otras ocasiones, el efecto provocado ante el mismo fenómeno era el contrario, un exceso de personal en los locales desocupado y sin comensales que atender. Ante esta falta de previsión y mala gestión del personal dentro de una misma cadena de restaurantes se planteó la posibilidad de elaborar una herramienta que pudiera tener en cuenta factores externos e internos a la hora de predecir el número de comensales que acudirían a cada local a lo largo del día. Esta estimación haría posible una mejor gestión de recursos y personal, pudiendo optimizar el número de empleados y distribución de los mismos en los distintos locales.

En este contexto fui contratada por la firma Price Waterhouse Coopers para desarrollar este proyecto como trabajo de fin de máster durante el periodo de prácticas de seis meses.

3.2 Motivación

La misión principal de las nuevas tecnologías es la de optimizar procesos y recursos para obtener mejores resultados a través de los datos. Tal y como se observa tras haber analizado la situación actual del sector de restauración, existen procesos que están lejos de ser óptimos. Entre ellos, la gestión del personal y la aplicación de herramientas de previsión para ajustar los procesos a las necesidades de los establecimientos aumentando la productividad y rentabilidad de los mismos aún está en pleno de desarrollo. Aunque existen programas destinados a automatizar la tarea de planificación o que ayudan a calcular el personal necesario, prácticamente ninguna herramienta a día de hoy realiza una previsión de afluencia en los establecimientos, teniendo en cuenta factores reales que afecten a las visitas, para poder realizar posteriormente una buena planificación. De hecho, las soluciones emergentes que afirman aplicar la analítica predictiva y algoritmos de inteligencia artificial aún deben demostrar la eficacia de las mismas y la realidad que se esconde tras de ellas.

Este proyecto nace de la necesidad de demostrar la posibilidad de aplicación de las herramientas de explotación de datos al sector de restauración con el objetivo de prever la cantidad de visitas que van a producirse en un restaurante. Esto permitirá poder planificar de forma óptima el personal necesario y los turnos más adecuados para ofrecer mejores servicios y obtener mayores beneficios. En esta previsión se tendrán en cuenta no solo el histórico de datos del propio establecimiento, si no también factores externos como el tiempo atmosférico, eventos sociales, situación geográfica, ofertas y promociones, índice de consumo, situación laboral...

La diferencia con el resto de soluciones existentes y en particular con aquellas que dicen aplicar la Inteligencia Artificial y el Machine Learning es que este proyecto tiene como objetivo realizar una solución completamente adaptada al cliente, desde su inicio. Para ello, se partirá de una base

de datos prototipo de un cliente para en primer lugar detectar según el histórico de visitas qué factores han podido influir en este hecho. Una vez detectados los factores se incluirán en un modelo de predicción que calculará las previsiones a corto y medio plazo del restaurante. Finalmente, se calculará el número óptimo de empleados que será necesario en cada momento para satisfacer la demanda cumpliendo las exigencias de satisfacción del restaurante. De modo que, en vez de adaptar el cliente a la herramienta se adaptará la herramienta al cliente, ofreciendo así una solución única adaptada a las necesidades.

Se pretende demostrar con este proyecto la importancia y el impacto que puede llegar a tener una buena gestión de personal en un restaurante. Los empleados de un restaurante son uno de los principales recursos que tienen para mejorar la experiencia del cliente en el negocio, así como para alcanzar la satisfacción del mismo. La mayoría de las empresas desconocen el impacto negativo que puede conllevar la falta o exceso de personal en momentos concretos del día. Ello puede verse traducido en enormes pérdidas potenciales, que al no estar cuantificadas ni medidas se tiende a no tenerlas en cuenta, lo cual supone un grave error. A lo largo del proyecto se analizará y medirá el impacto de la aplicación de las nuevas tecnologías a la planificación y gestión del personal, para demostrar así la eficacia y conveniencia del proyecto.

Los siguientes datos extraídos de diversos informes (White House Office of Consumer Affairs) y fuentes (McKinsey, Bain...), ayudan a demostrar la importancia de mantener un buen servicio al cliente:

- Por cada queja de un cliente, existen otros 26 clientes no satisfechos que se han mantenido en silencio.
- El 96% de los clientes no satisfechos no ponen quejas. Sin embargo, el 91% de ellos simplemente no regresarán.
- Un cliente no satisfecho contará sus experiencias a 9 – 15 personas. Alrededor del 13% de los clientes no satisfechos se lo manifestarán a más de 20 personas.
- Clientes satisfechos que han resuelto sus problemas contarán su experiencia a 4 – 6 personas.
- Hacen falta 12 experiencias positivas para recuperar una mala experiencia sin resolver.
- Una reducción del 5% en la proporción de errores con clientes puede aumentar los beneficios entre un 5 - 95 %.
- Adquirir un nuevo cliente cuesta 6-7 veces más que mantener a un cliente existente.

3.3 Objetivos del proyecto

Los principales objetivos que se pretenden llevar a la cabo a lo largo del proyecto son los siguientes:

- Analizar las metodologías de gestión de personal que se están llevando a cabo en la actualidad con el objetivo de identificar puntos de mejora.
- Detectar los factores que influyen en la afluencia en restaurantes y desarrollar una simulación de negocio y base de datos realista que sirva de prototipo para el proyecto, incluyendo dichos factores y comportamientos reales.
- Implementar herramientas de análisis de Machine Learning al modelo para estudiar los datos generados e identificar posibles patrones de comportamiento de clientes de una cadena de restauración.
- Analizar y validar distintos modelos de predicción basados en técnicas de Inteligencia Artificial en función de los objetivos buscados y los resultados obtenidos.

- Configuración y parametrización de un modelo de predicción de afluencia teniendo en cuenta los factores influyentes obtenidos, aplicando el modelo a la base de datos existente. Posterior verificación de su eficacia y precisión.
- Utilizar los resultados del modelo de predicción para optimizar la gestión y planificación del personal y del stock dentro del restaurante y analizar los resultados.
- Aplicación de la metodología estándar en la construcción de una herramienta con Machine Learning.
- Determinar el impacto económico de la implantación del modelo de predicción propuesto y la optimización del número de empleados.
- Estudiar el impacto a nivel organización del personal de la implantación de la herramienta.
- Desarrollar una herramienta que permita ser utilizada por un cliente real, de forma que al introducir su base de datos pueda visualizar en la herramienta la predicción de comensales y la optimización del personal.

3.4 Metodología y recursos a emplear

El objetivo principal del proyecto es el desarrollo de una herramienta que permita incorporar la base de datos de un cliente real a la misma y que dicha herramienta sea capaz de a partir de esos datos del histórico realizar una predicción de afluencia teniendo en cuenta diversos factores y calcular el número de empleados óptimo para cada turno.

Debido a la dificultad de acceso a una base de datos real de un cliente del sector de restauración, es necesaria la simulación de un histórico de afluencia que servirá como base de datos de una cadena de restauración. Se pretenderá que dicha simulación sea lo más realista posible y que refleje la realidad con un nivel alto de perfección. Para ello se empleará el método Delphi, método que consiste en la estructuración de un proceso de comunicación grupal que permite a un grupo de individuos tratar un problema complejo, en este caso la predicción de afluencia en una cadena de restauración para elaborar la base de datos que se utilizará para el modelo de predicción como histórico. Este método se basa en la utilización sistemática de un juicio intuitivo emitido por un grupo de expertos, que permite conocer con una mayor seguridad cómo va a evolucionar algún elemento o comportamiento, que en este caso son las visitas a los distintos locales de una cadena de restauración, mediante una serie de técnicas de prospección. El proceso consiste en reunir a un grupo de expertos a los que se somete a un cuestionario y tras analizar los resultados se entrega de nuevo el cuestionario a los expertos y se les pide que vuelvan a responder. El objetivo es llegar a un consenso entre los cuestionarios recibidos. En este proceso se analizarán cuáles son los factores que afectan e influyen a la afluencia en restaurantes, para poder incluirlos en la simulación.

Este base de datos tiene la función de prueba de concepto, ya que el objetivo es que cualquier cliente pueda incorporar fácilmente su base de datos en la herramienta que se ha desarrollado en este proyecto.

La siguiente etapa corresponde al desarrollo del modelo de predicción, incorporando los factores extraídos de la prueba de concepto al propio modelo. Para ello se utilizarán diversos algoritmos, aplicando herramientas de Machine Learning, con el objetivo de implementar aquel con el que obtengamos mejores resultados y precisión. Para elaborar el modelo predictivo se llevará a cabo el proceso mostrado a continuación:

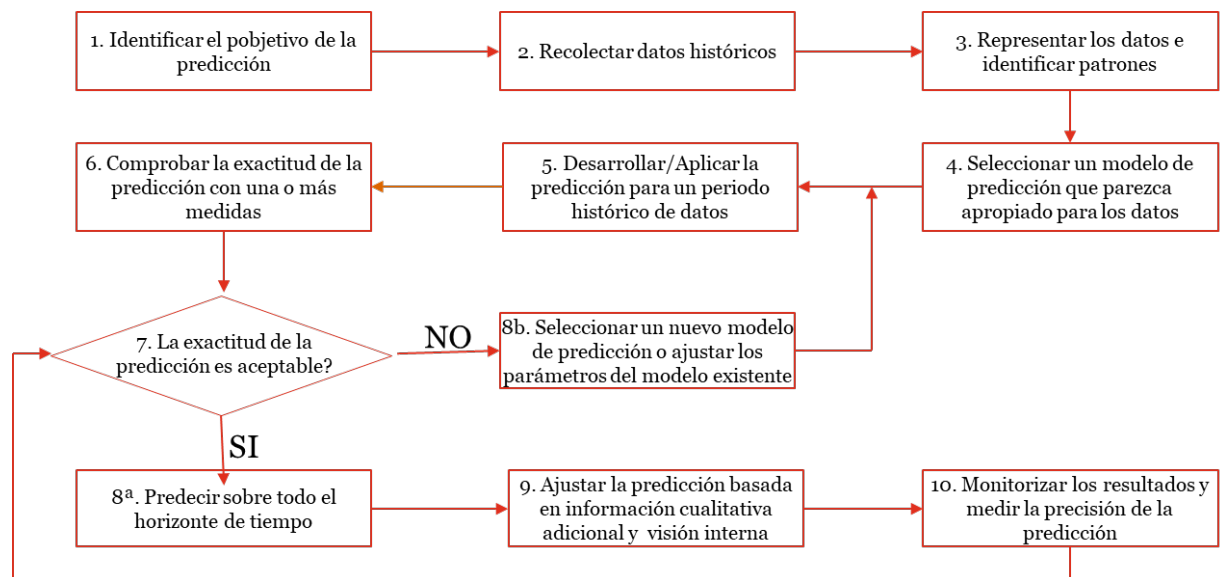


Figura 7. Proceso de elaboración de un modelo de predicción

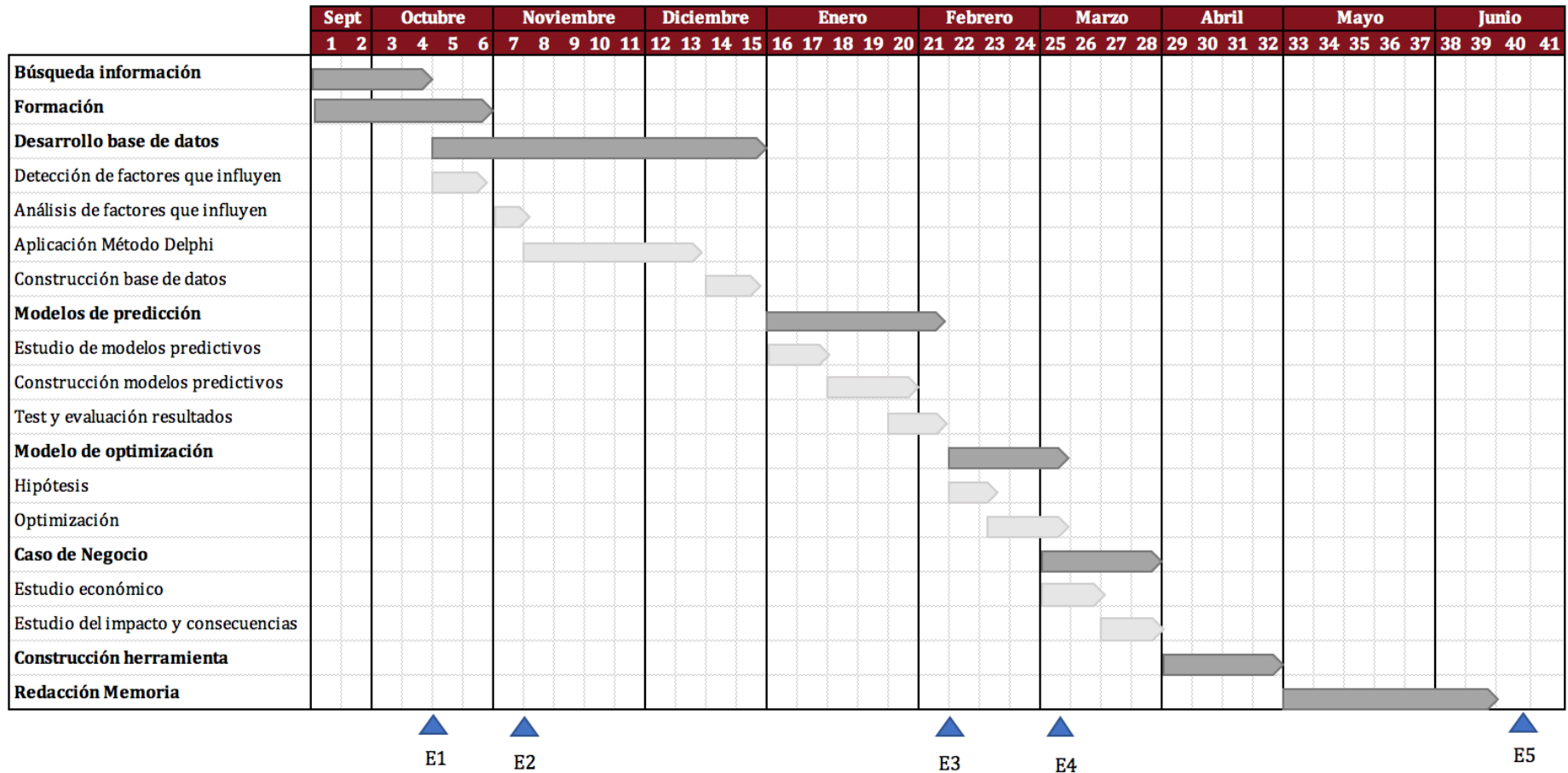
Tanto la prueba de concepto como el desarrollo del modelo de predicción se desarrollarán en el entorno de programación Python, lo cual requerirá un periodo de formación y aprendizaje.

La predicción de afluencia permitirá no solo ajustar la planificación del personal a las necesidades, también hará posible la optimización del cálculo de demanda y mejorará el sistema de pedidos. La herramienta desarrollada incluirá un modelo de optimización que permita calcular el número óptimo de empleados que se necesita en cada momento ajustándolo al nivel de exigencia de atención al cliente requerido por la cadena de restauración.

Por último, se llevará a cabo un estudio económico de la solución propuesta, así como un análisis de costes para distintos escenarios donde se podrá demostrar el impacto de la utilización de dicha herramienta.

3.5 Planificación

A continuación se adjunta un cronograma con la programación del proyecto, llevado a cabo junto con el director del mismo.



HITOS	
E1	Entrega Anexo A
E2	Entrega Anexo B
E3	Presentación Anexo B
E4	Presentación estado actual TFM
E5	Presentación final TFM

CAPÍTULO IV : SOLUCIÓN DESARROLLADA



4. Solución desarrollada

4.1 Desarrollo del modelo de simulación de afluencia

Simular consiste en obtener resultados futuros para una variable aleatoria a partir del comportamiento del pasado de esa misma variable aleatoria, asumiendo que las propiedades estocásticas se mantendrán en el futuro. Para llevar a cabo la simulación habrá que establecer una serie de hipótesis que posteriormente deberán ser contrastadas y verificadas. Ese proceso de contraste y verificación se realizará con un grupo de expertos en el sector a través del método Delphi, mencionado en la sección anterior.

Según R.E. Shannon¹, “La simulación es el proceso de diseñar un modelo de un sistema real y llevar a término experiencias con él, con la finalidad de comprender el comportamiento del sistema o evaluar nuevas estrategias – dentro de los límites impuestos por un cierto criterio o un conjunto de ellos – para el funcionamiento del sistema”. Este es en particular el objetivo de este apartado, realizar una simulación de cuál y cómo sería la afluencia de comensales en una serie de restaurantes pertenecientes a una cadena de restauración. Con este apartado se pretende aprender a simular el funcionamiento de un negocio de restauración para poder evaluar el mismo y a la vez asesorar qué medidas sería conveniente aplicar para mejorar su rendimiento bajo ciertas condiciones.

4.1.1 Introducción

Puesto que el objetivo del proyecto es realizar una predicción de afluencia, debemos contar con una serie de datos de visitas a restaurantes para poder introducir al modelo de predicción como histórico. Debido a que no siempre es posible acceder a esos datos se van a utilizar los resultados de la simulación del negocio de restauración como histórico que introduciremos posteriormente al modelo predictivo. Es por ello que es muy importante que la simulación se ajuste lo más posible a la realidad.

En este apartado se explicará el desarrollo del modelo de simulación y se analizarán los resultados obtenidos para verificar la viabilidad del mismo. El modelo de simulación se va a desarrollar en el entorno de programación Python. A continuación se detallarán los datos de entrada necesarios para la ejecución.

Se pretende realizar un modelo de simulación de afluencia para una cadena de restaurantes con distintas características. Para realizarlo es preciso conocer los patrones de comportamiento de los comensales y los factores que pueden influir en el nivel de afluencia en restauración según distintos parámetros. Para ello, es necesario también conocer las características de los restaurantes que vamos a simular. En este caso se va a llevar a cabo un modelo simplificado de una cadena de restaurantes, de manera que posteriormente se puedan escalar los resultados si se desea. Dicha cadena contará con dos tipos de restaurantes, uno de comida rápida y otro de comida tradicional. Cada tipo de restaurante contará a su vez con tres locales, que estarán situados en tres tipos de zonas para ambos casos.

La primera zona consiste en un centro comercial, la segunda en un barrio residencial y la última en una zona de oficinas. Se han escogido estas zonas ya que se ha considerado una diferencia significativa en el comportamiento de los comensales en cada una de ellas.

En la siguiente tabla se pueden resumir las características de cada uno de los restaurantes con los que vamos a contar para la simulación:

¹ Shannon, Robert; Johannes, James D. (1976). «[Systems simulation: the art and science](#)». *IEEE Transactions on Systems, Man and Cybernetics*. 6(10). pp. 723-724.

Nombre	Tipo	Ubicación	Inicio servicio	Fin servicio	Terraza	Parking	Capacidad máxima
Restaurante 1	comida rapida	centro comercial	13:00	17:00	no	si	150
Restaurante 2	comida rapida	residencial	13:00	16:00	si	no	150
Restaurante 3	comida rapida	oficina	12:00	16:00	no	no	150
Restaurante 4	comida tradicional	centro comercial	13:00	17:00	si	si	80
Restaurante 5	comida tradicional	residencial	13:00	17:00	no	si	80
Restaurante 6	comida tradicional	oficina	12:00	16:00	si	no	80

Tabla 1. Caracterización de restaurantes

Como se puede observar, el único servicio que se va a simular es el de la comida, siendo el horario distinto para cada restaurante en función de la ubicación del mismo. Por otro lado, cada local tiene distintas peculiaridades. Por ejemplo, ciertos restaurantes disponen de terraza y/o parking. Por último, la capacidad máxima de los restaurantes de comida rápida es de 150 comensales y la de los restaurantes de comida tradicional, 80, entendiendo que el tamaño del local y su capacidad es la misma para cada uno de los locales dentro del mismo tipo de comida.

El hecho de caracterizar los restaurantes para los cuáles se va a realizar una simulación es importante puesto que estas características tendrán una influencia respecto al número de personas que acudirán al restaurante. Por ejemplo, si se trata de un restaurante que dispone de parking accesible para clientes habrá más posibilidades de que ante un día lluvioso, los comensales prefieran acudir a dicho restaurante que a uno que no disponga de parking.

Otro de los datos de entrada para realizar la simulación será el calendario. Es preciso conocer para qué fechas se va a necesitar realizar la simulación y qué ocurrió en esas fechas. De esta forma es importante conocer factores como: temperatura media, precipitaciones, día de la semana, mes, año,... Otro dato importante a tener en cuenta es el hecho de que el día sea festivo o no. Esto puede afectar de manera positiva o negativa a la afluencia en restaurantes, en función de la ubicación del mismo.

A modo de resumen se presenta a continuación una tabla ejemplo de los datos de entrada de calendario para el modelo. En este caso el calendario de simulación partía del 1 de Enero del 2018 y finalizaba el 28 de Febrero de 2019.

Fecha	Día	Tmed	Precip	Festivo	Festivo escolar navidades
01/01/2018	lunes	7.4	0	1	0
02/01/2018	martes	10.4	0	0	1
03/01/2018	miércoles	10.8	0	0	1
04/01/2018	jueves	10.9	0	0	1
05/01/2018	viernes	8.8	6	0	1
06/01/2018	sábado	4.4	18.6	1	0
07/01/2018	domingo	2.9	5.8	0	0
08/01/2018	lunes	4.4	0.1	0	0
09/01/2018	martes	1.9	1.6	0	0
10/01/2018	miércoles	5.2	0.6	0	0

Tabla 2. Caracterización de calendario simulación

Las unidades de medida de la temperatura media son grados centígrados (°C) y de las precipitaciones son milímetros. Se mide la altura de la lámina de agua recogida en una superficie plana. 1 milímetro de agua de lluvia equivale a 1 litro de agua por m².

4.1.2 Reglas de generación

Para realizar la simulación ha sido necesario establecer una serie de reglas de comportamiento que reflejaran la realidad de la forma más aproximada. Para ello ha sido necesario un estudio previo gracias a la información aportada por consultores y profesionales del sector, y el análisis de ciertas bases de datos existentes en la firma PwC que han hecho posible llegar a las conclusiones que se presentan a continuación. Este método iterativo de consulta y verificación de resultados con un grupo de expertos se conoce también como Método Delphi. Se trata de un proceso largo y costoso por el número de iteraciones y de personas implicadas en él.

El procedimiento consiste en generar una serie de factores que reflejen la probabilidad de que un restaurante tenga cierta ocupación un día en particular. Posteriormente este factor será multiplicado por el nivel de afluencia media de cada restaurante. En este caso se ha establecido también a través del Método Delphi, junto con el grupo de expertos, que el nivel de afluencia media es de un 65% de la capacidad máxima.

Hasta llegar al factor final que se utilizará para la simulación, se han realizado numerosas iteraciones en las que en cada caso se analizaban los resultados obtenidos de afluencia y junto con el grupo de expertos se decidían las modificaciones y nuevas adaptaciones de los factores. Este proceso se repitió durante tres meses hasta que se alcanzaron los resultados esperados y se concluyó que el resultado de la simulación era realista y no se podía mejorar más.

De acuerdo a los factores que tienen una influencia directa en el número de visitantes que acuden a un restaurante, se han establecido seis reglas distintas. Las dos primeras son reglas aditivas mientras que las cuatro últimas son multiplicativas. Es decir, partiendo del valor inicial “1”, se irán sumando o restando valores en función de los resultados de las dos primeras reglas, y a continuación, se multiplicará el resultado por otros factores derivados de la regla número 3,4,5 y 6.

Para el calendario de simulación definido se irán recorriendo cada una de las normas para cada uno de los días en función de las características que presenten los mismos, obteniendo así el factor asociado a cada uno de los días. Las reglas establecidas son las siguientes:

Regla 1: Regla diaria

En función del día de la semana en el que nos encontremos y de la ubicación del propio restaurante se sumará o restará un factor al valor inicial. Hay que tener en cuenta que estos factores se asocian al turno de comida. En la tabla que se muestra a continuación se pueden observar dichos factores.

Centro Comercial		Barrio Residencial		Zona Oficinas	
Lunes	-0.4	Lunes	-0.6	Lunes	0.4
Martes	-0.35	Martes	-0.5	Martes	0.3
Miércoles	-0.35	Miércoles	-0.4	Miércoles	0.2
Jueves	-0.15	Jueves	0.1	Jueves	0.3
Viernes	0.4	Viernes	0.3	Viernes	0.1
Sábado	0.5	Sábado	0.7	Sábado	-0.5
Domingo	0.2	Domingo	0.5	Domingo	-0.5

Tabla 3. Factores regla diaria

La tendencia general en restaurantes ubicados tanto en centros comerciales como en barrios residenciales es que a principios de semana la afluencia sea menor que a finales. A partir de los

viernes las visitas aumentan, debido a que la gente no trabaja y puede disfrutar de más tiempo libre. No obstante, en las zonas de oficina ocurre lo contrario. A lo largo de la semana la afluencia se mantiene en unos niveles similares, mientras que el fin de semana disminuye drásticamente al reducirse en gran medida el flujo de gente en la zona.

Regla 2: Regla mensual

En este caso, el momento del mes en el que se encuentre un cliente afecta también a la decisión de acudir a un restaurante o no. Por ejemplo, a principios de mes, momento generalmente próximo a la recepción del salario, la gente es más propensa a realizar más gasto. Mientras que, a finales de mes, la gente suele consumir menos. De modo que los factores que se deben aplicar al resultado obtenido tras la regla 1 son los siguientes:

Para todo tipo de local	
Días 1-8	0.07
Días 9-15	0.05
Días 15-19	0
Días 20-24	-0.05
Días 25-31	-0.07

Tabla 4. Factores regla mensual

Regla 3: Regla eventos o festivos

En los días festivos se producen distintas situaciones según dónde esté localizado el restaurante. Por ejemplo, al tratarse de un día festivo, la gente no acudirá a la oficina, por lo que las visitas en restaurantes situados en zonas de oficinas se verán disminuidas. Sin embargo, en barrios residenciales o centros comerciales aumentará bastante el número de los mismos ya que es más probable que la gente aproveche los días festivos para realizar actividades fuera de casa. Al mismo tiempo, al ofrecer los centros comerciales la posibilidad de realizar compras y otras actividades de ocio, es ligeramente más probable que acuda más gente en día de festivo a centros comerciales.

Estos factores se multiplicarán por el resultado obtenido tras la haber aplicado la segunda regla. A continuación, el resultado pasará a la siguiente regla como dato de entrada.

Centro Comercial		Barrio Residencial		Zona oficinas	
Sí	*1.3	Sí	*1.25	Sí	*0.6

Tabla 5. Factores regla eventos y festivos

Regla 4: Regla climatológica

Respecto al clima, existen distintas posibilidades que se combinan con las características de los propios restaurantes. De ellas, las posibilidades que más influencia pueden tener son:

- Que haya precipitaciones y el restaurante tenga parking
- Que haya precipitaciones y el restaurante no tenga parking
- Que no haya precipitaciones, el restaurante tenga terraza y que la temperatura media sea superior a 20°C

- Que no haya precipitaciones, el restaurante no tenga terraza y que la temperatura media sea superior a 20°C.

Para que se considere el fenómeno de precipitación como influyente en la afluencia es preciso que haya llovido más de 10 mm, ya que, en caso contrario, las pequeñas precipitaciones pueden no tener un impacto directo en el comportamiento de los clientes.

Estos cuatro fenómenos son los que tendrán más impacto a la hora de simular la afluencia en restaurantes. En función de si se trata de un restaurante ubicado en un centro comercial, un barrio residencial o una zona de oficinas habrá que multiplicar el resultado obtenido después de la regla 3 por los siguientes factores:

	Centro Comercial		Barrio Residencial		Zona Oficinas	
	Festivo	No Festivo	Festivo	No Festivo	Festivo	No Festivo
Lluvia + Parking	*1.25	*1.1	*1.25	*1.15	*1	*1.15
Lluvia + No Parking	*0.85	*0.7	*0.8	*0.65	*1	*0.9
No lluvia + terraza + Tmed>20	*1.2	*1.1	*1.3	*1.15	*1	*1.3
No lluvia + No terraza + Tmed>20	*1	*1	*1	*1	*1	*1.15

Tabla 6. Factores regla climatológica

Tal y como se observa en la tabla anterior, cuando un restaurante tiene parking y está lloviendo en el exterior, aumenta la posibilidad de que los comensales acudan a dicho local. Del mismo modo, el hecho de no tener parking hace que se reduzcan las probabilidades. Por otro lado, si no se producen precipitaciones y la temperatura ambiente media es superior a 20°C, las probabilidades de que haya más visitas en un restaurante de cualquier tipo aumentan.

Regla 5: Regla estacionalidad

Durante los meses de verano la actividad de los restaurantes situados en ciudades como Madrid, como en este caso, se ve reducida debido al periodo vacacional. Durante los meses de Julio y Agosto, coincidiendo con las vacaciones escolares y de la mayoría de trabajadores, se reduce en gran parte el número de habitantes en ciudades como Madrid. Por lo tanto, es lógico esperar que el número de visitas se reduzca de manera considerable.

	Centro Comercial	Barrio Residencial	Zona oficinas
Julio	*0.75	*0.75	*0.75
Agosto	*0.65	*0.5	*0.6

Tabla 7. Factores de regla estacionalidad

Regla 6: Regla vacaciones escolares

En este caso, se han considerado de manera independiente las vacaciones escolares del periodo navideño ya que tienen un efecto contrario al producido en las vacaciones de verano. Durante este periodo se ha observado que el nivel de afluencia en restaurantes aumenta en zonas de centros comerciales y barrios residenciales y disminuye en zonas de oficina. Ello es debido a las numerosas actividades y reuniones navideñas, períodos de alta frecuencia de compras... Los factores más apropiados que se aplicarán como consecuencia de esta regla se muestran en la siguiente tabla:

	Centro Comercial	Barrio Residencial	Zona oficinas
Festivo Escolar	*1.2	*1.1	*0.8
No Festivo Escolar	*1	*1	*1

Tabla 8. Factores de regla vacaciones escolares

Finalmente, una vez tenemos el factor final obtenido tras la aplicación consecutiva de las seis reglas, se debe añadir una variabilidad aleatoria. Para ello, multiplicaremos el factor final por un número aleatorio entre 0.8-1.2. De esta forma, estaremos añadiendo una variabilidad aleatoria del $\pm 20\%$ al resultado final del valor del factor de afluencia.

Una vez se ha obtenido el factor final, se multiplicará dicho valor por el promedio de afluencia de cada restaurante como se ha indicado anteriormente. En este caso, y tras haber analizado información acerca del sector, se ha estimado que la afluencia promedio corresponde a un 65% de la capacidad máxima de cada restaurante.

4.1.3 Resultados obtenidos

La simulación de afluencia anual para los 3 primeros restaurantes, caracterizados por servir comida rápida es la siguiente:

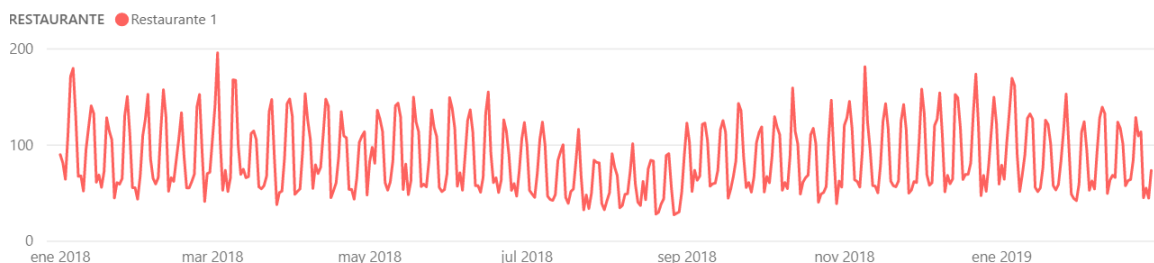


Figura 8. Simulación afluencia Restaurante 1



Figura 9. Simulación afluencia Restaurante 2

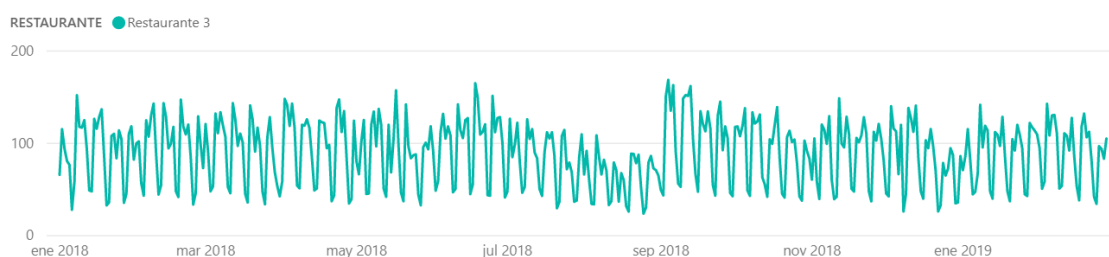


Figura 10. Simulación afluencia Restaurante 3

La simulación anual para los segundos tres restaurantes, que ofrecen un servicio de comida tradicional es:



Figura 11. Simulación afluencia Restaurante 4

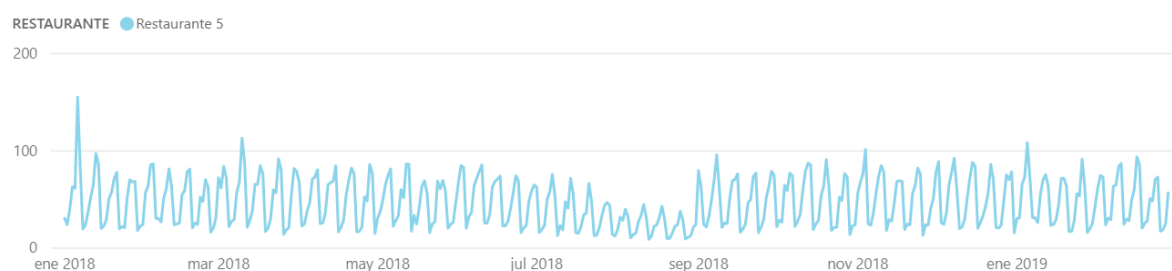


Figura 12. Simulación afluencia Restaurante 5

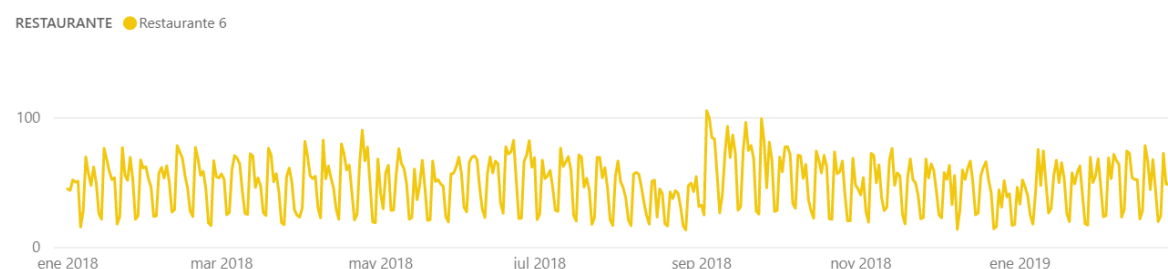


Figura 13. Simulación afluencia Restaurante 6

Se observa en la simulación presentada cierta estacionalidad anual, disminuyendo el número de visitas en restaurantes en los meses de verano, y presentando ciertos picos en fechas señaladas que coinciden con festivos o periodos de lluvia o condiciones propensas al incremento de afluencia. Además, se observa como en ciertas ocasiones se sobrepasa el máximo de la capacidad de los locales, lo que simulará situaciones en las que se generarán colas en dichos locales y tiempos de espera para ser atendidos, llegando incluso a no poder satisfacer toda la demanda.

En la siguiente gráfica se puede observar la distribución del promedio de las visitas según el día de la semana. Es importante recordar que el Restaurante 1 y Restaurante 4 están situados en centros comerciales, el Restaurante 2 y 5 en zonas residenciales y el 3 y 6 en zonas de oficina. Este factor es importante para entender la distribución semanal.

Para los restaurantes situados en zonas residenciales y en zonas de centros comerciales las visitas aumentan a partir del viernes. El fin de semana estos locales registran un aumento considerable de visitas mientras que los restaurantes de oficinas experimentan el efecto contrario. Durante la semana, el registro de visitas es inferior para los restaurantes situados en centros comerciales y barrios residenciales que en aquellos situados en zonas de oficina. Es por ello que se considera que la localización de los restaurantes de una cadena tiene especial importancia a la hora de estimar el número de visitas que se van a producir.

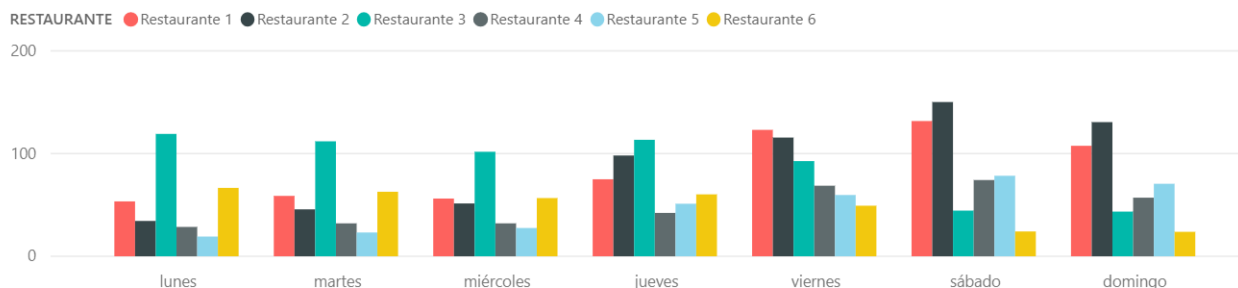


Figura 14. Promedio de afluencia semanal

Por otro lado, es posible conocer también el promedio de afluencia para cada mes del año. La diferencia de afluencia observada entre los tres primeros restaurantes y los tres segundos es debida a que la capacidad máxima de los primeros es de 150 comensales mientras que la de los tres segundos es 80 comensales. Observamos como la distribución mensual a lo largo del año es similar para los restaurantes, registrando menores ventas en los meses de julio y agosto y aumentando la afluencia durante festivos y acontecimientos favorables.

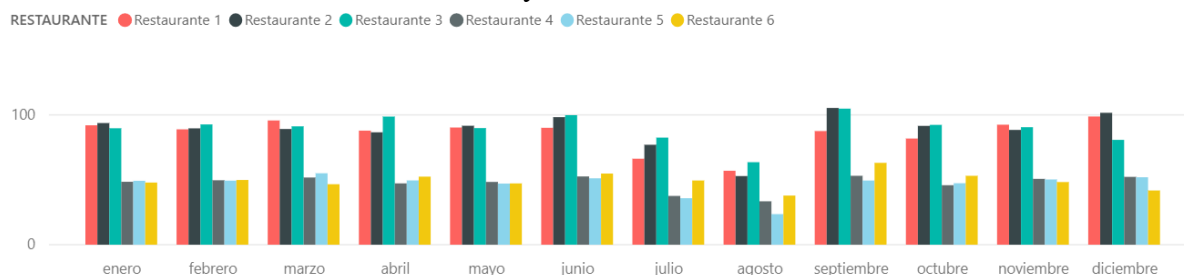


Figura 15. Promedio de afluencia mensual

La variación del promedio de afluencia de cada tipo de restaurante se puede observar mejor en las Figura 16 y Figura 17.



Figura 16. Promedio de afluencia en restaurantes de tipo "Comida rápida"

Se puede observar que para el Restaurante 1 (situado en centro comercial) los meses de mayor afluencia son marzo, noviembre y diciembre y los de menor afluencia los meses de julio, agosto y octubre. Para el Restaurante 2 (zona residencial) se producen más visitas en los meses de junio, septiembre y diciembre, y menos visitas en los meses de julio, agosto y abril. En cambio, en el Restaurante 3 (zona de oficinas) el máximo se produce en septiembre, a continuación junio y abril. El número mínimo se registra en agosto, diciembre y noviembre. Observamos cómo en el mes de agosto, que coincide con el mes de menor actividad en Madrid debido al período vacacional se produce un importante descenso en el número de visitas para todos los restaurantes. En el período navideño aumentan las visitas en los restaurantes de zonas de ocio y residenciales debido a las actividades navideñas, y disminuyen en las zonas de oficinas como consecuencia del período vacacional.

A continuación, se muestra la variación del promedio de afluencia para los restaurantes 4,5 y 6. Del mismo modo que ocurre en el caso anterior, el mes de agosto es el mes que registra menos visitas para todos los restaurantes. En el mes de diciembre aumentan las visitas en centros comerciales y barrios residenciales, pero disminuyen en zonas de oficina. Para esta zona, el mes de septiembre es el que registra el máximo número de visitas en promedio, siendo junio y abril los siguientes. Marzo, Junio y Septiembre son los meses con mayor afluencia para los restaurantes 4 y 5.

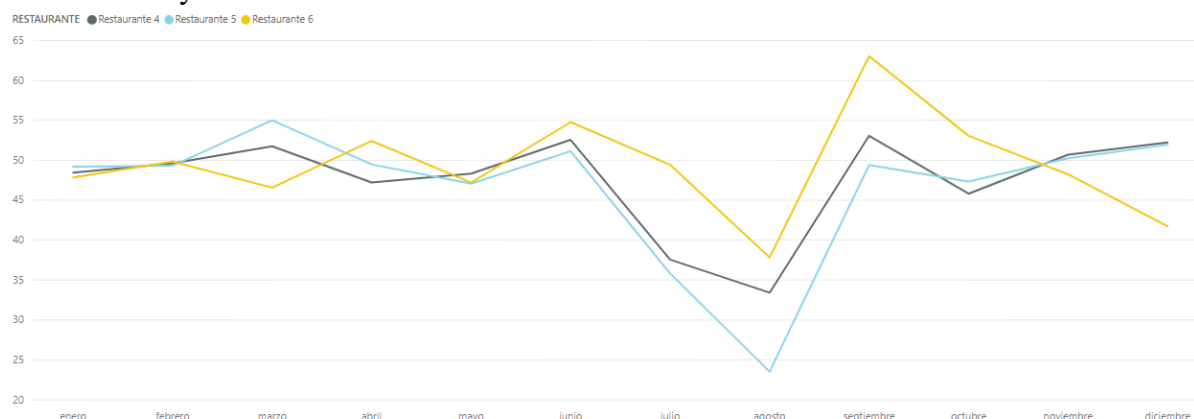


Figura 17. Promedio de afluencia en restaurantes de tipo "Comida tradicional"

4.1.4 Aplicación real

Esta simulación puede ser muy útil para dos tipos de clientes. Por un lado, puede ser útil para los gestores de cadenas de restauración o incluso de cualquier tipo de local de hostelería, sea del tamaño que sea. Esto le permitiría tener una visión de cómo va a evolucionar su negocio, y en particular cada local individualmente y poder actuar en consecuencia. Puede promover la creación de promociones para los meses de menor afluencia y ayudar a la previsión de stock para aquellos meses de mayor afluencia por ejemplo. Permite optimizar la gestión de cada centro independientemente, para mejorar el rendimiento de la cadena en términos globales.

También puede ser útil para aquellos emprendedores que estén pensando en montar un negocio de este tipo, y no conozcan bien el sector o simplemente quieran tener una mayor visión de la evolución de un posible local en una determinada localización, teniendo en cuenta diversos factores y situaciones.

Por otro lado, también podría ser útil para negocios como el de la consultoría ya que en muchas ocasiones la falta de información en el inicio del proyecto complica y alarga los estudios y análisis. El hecho de tener este tipo de modelo de simulación permitiría hacer un estudio claro y realista del potencial cliente en el sector de restauración desde un primer momento, permitiendo así adecuar mejor y más rápido las soluciones al cliente una vez aporten la información real.

4.2 Desarrollo del modelo de predicción optimizado

4.2.1 Introducción

En este apartado se desarrolla la parte principal del proyecto, la predicción de afluencia. Para ello se van a implementar distintas técnicas de Inteligencia Artificial haciendo uso de herramientas de *Machine Learning* o aprendizaje automático, que son una combinación de modelos matemáticos, estadísticos y probabilísticos capaces de predecir un resultado. Con estas técnicas se realizará una predicción de demanda en el sector de restauración teniendo en cuenta los datos históricos generados específicamente para ello. El objetivo será comparar la precisión de cada una de las técnicas utilizadas y utilizar en cada caso aquella que ofrezca resultados más fiables y precisos. Se pretende comprender el funcionamiento de los mismos así como la parametrización óptima de cada uno de ellos.

Los algoritmos de aprendizaje automático (Machine Learning) son los grandes pilares y retos de la transformación digital en el sector de restauración. El análisis predictivo permitirá adelantarse a las necesidades de los clientes, establecer medidas de actuación incorporando información en tiempo real, mejorar la gestión de cada local y ofrecer un servicio óptimo aumentando la satisfacción del cliente.

La metodología general y procedimiento estándar de construcción de un modelo de aprendizaje automático se resume en los siguientes pasos: recolección de datos, separación del conjunto de datos en set de entrenamiento y set de testeo o validación, aprendizaje del algoritmo, evaluación y por último, optimización o refinamiento. Para optimizar los modelos habrá que realizar un ajuste de sus parámetros sin llegar a sobre-ajustar, para reducir los errores obtenidos al máximo, consiguiendo así una predicción lo más precisa posible. En la siguiente imagen se puede observar cuál es el procedimiento mencionado:

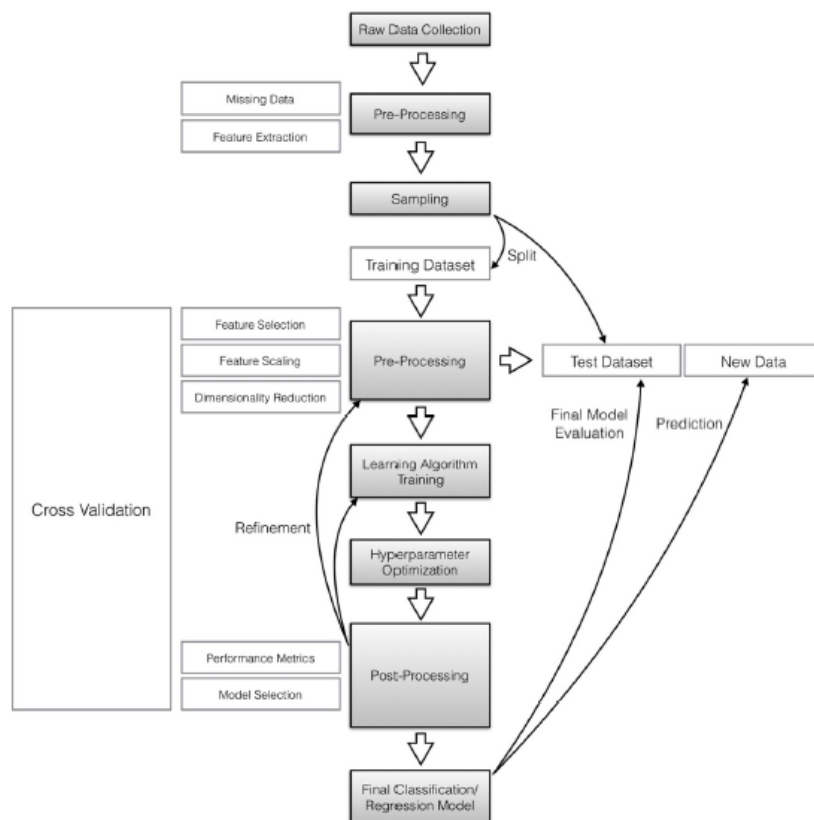


Figura 18. Procedimiento estándar de construcción de un modelo de aprendizaje automático

4.2.2 Modelos de predicción

El modelo de predicción de afluencia permitirá conocer en base a información pasada sobre las visitas realizadas a dicho local, junto con las condiciones atmosféricas, geográficas y demográficas presentes, pasadas y futuras, el número total de comensales que acudirán al local en cada momento.

Las ventajas que aporta la implementación de un modelo predicción para estimar la afluencia en restaurantes dentro de una cadena de restauración son:

- Permite ajustar el nivel de personal necesario en cada local, evitando exceso y falta de personal en momentos clave del día y así ofrecer un servicio óptimo en cada caso.
- Mejora la gestión del stock y de materia prima.
- Permite estimar los momentos de menor afluencia y elaborar promociones personalizadas para contrarrestar este efecto.
- Permite detectar los picos de afluencia y elaborar medidas para evitar colas y agilizar el servicio.
- Permite elaborar turnos equilibrados que aseguren una buena calidad de trabajo para los empleados que cumpla con las regulaciones laborales y que estimule a los empleados a seguir trabajando en la empresa.

Para los modelos de predicción que se presentan a continuación se utilizará la simulación de afluencia realizada en el apartado anterior como histórico.

Es importante evaluar la precisión de la predicción. La precisión del modelo solo se puede determinar considerando cómo de bien el modelo actúa sobre un nuevo conjunto de datos que no fueron utilizados para ajustar y entrenar el modelo. Para calcular la medida de precisión, se obtendrá previamente el valor del MAPE (*Mean absolute percentage error*), valor que mide tamaño del error absoluto del conjunto de predicciones de datos en términos de porcentaje. Para reflejar la precisión del modelo se restará este valor al total del 100%. De esta forma obtendremos el porcentaje de exactitud con el que el modelo actúa sobre el conjunto de datos y la calidad de predicción que ofrece.

Para ello se va a dividir el conjunto de datos del histórico que poseemos (13 meses en total) en dos partes. La primera corresponderá con el conjunto de muestreo con el que se entrenará el modelo, y el segundo conjunto será el de testeo. El conjunto de datos de entrenamiento se usa para estimar los parámetros del método de predicción. El conjunto de datos de testeo será utilizado para calcular la precisión y eficacia del modelo. En este caso, se ha utilizado un periodo de muestreo de 12 meses (Enero 2018 – Enero 2019) y un periodo de testeo de 1 mes para el modelo (Febrero 2019). Posteriormente se realizará la predicción real sobre el mes siguiente al del periodo de testeo (Marzo 2019).

Los algoritmos de Aprendizaje Supervisado (Machine Learning) que se han utilizados son:

4.2.2.1 Modelo Random Forest

El algoritmo Random Forest se trata de un algoritmo de aprendizaje supervisado capaz de realizar una predicción añadiendo un factor de aleatoriedad al mismo tiempo. Es una herramienta de modelización predictiva y no una herramienta descriptiva. El modelo de predicción funciona de forma que crea una serie de árboles de decisión y los une con el objetivo de obtener una predicción más precisa y estable. Está basado en el principio de que una combinación de modelos de aprendizaje mejora el resultado general de la predicción.

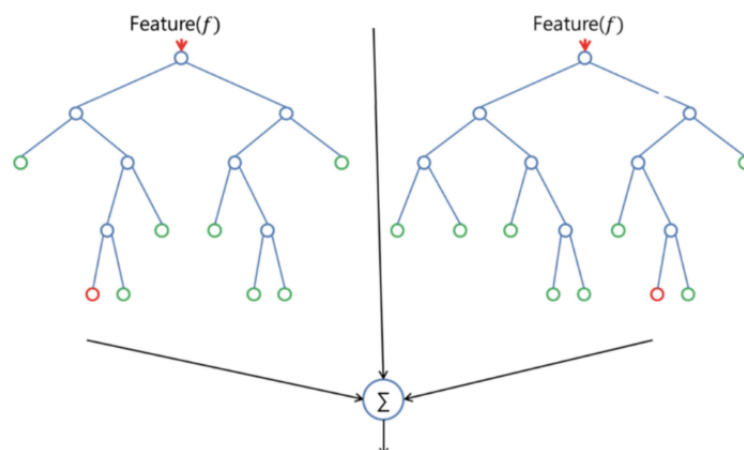


Figura 19. Árboles de decisión en Random Forest

Una de las ventajas que posee el modelo Random Forest frente a otros modelos es que puede ser utilizado tanto para problemas de clasificación como problemas de regresión.

En vez de buscar la característica más importante al dividir cada nodo, lo que este algoritmo busca es la mejor de las características existentes dentro de un subconjunto aleatorio de características. Es decir, para dividir cada uno de los nodos el algoritmo solo tiene en cuenta un subconjunto aleatorio de características. Además, para hacer que los árboles de decisión sean más aleatorios este modelo usa una serie de umbrales aleatorios adicionales para cada característica en lugar de buscar los mejores umbrales posibles como suelen hacer los árboles de decisión normales.

Otra de las ventajas que tiene este algoritmo es el hecho de que es muy sencillo medir la importancia relativa que tiene cada característica en la predicción. El entorno de programación Python posee una librería denominada Sklearn extremadamente útil para realizar esta función. Esta librería mide la importancia de una característica observando en qué medida los nodos de los árboles de decisión que usan dicha característica reducen la impureza en todos los árboles del “bosque”. El algoritmo calcula automáticamente esta puntuación para cada característica después de haber realizado un entrenamiento y posteriormente escala los resultados, de modo que la suma de todas las importancias sea uno. Una vez la importancia de cada característica es conocida, es posible tomar la decisión de qué características debemos eliminar debido a que no contribuyen lo suficiente al proceso de predicción. Una de las reglas generales en Machine Learning es que cuantas más características posea un sistema más probable será que el modelo sufra sobreajustes. Se pueden observar las importancias obtenidas para cada variable de cada restaurante en el *Anexo C – Importancia de variables según Random Forest*.

La diferencia principal de este algoritmo con los árboles de decisión generales se basa en que los árboles de decisión utilizan un set de datos de entrenamiento con una serie de características y etiquetas para generar una serie de reglas que serán utilizadas para realizar la predicción. No obstante, el algoritmo Random Forest selecciona observaciones y características de forma aleatoria y construye una serie de árboles de decisión con la finalidad de realizar un promedio de los resultados. Por otro lado, los árboles de decisión más “profundos”, a diferencia del algoritmo Random Forest, puede sufrir sobreajustes. Random Forest está preparado generalmente para evitar los excesos de adaptación. Para ello crea subconjuntos aleatorios de las características y construye árboles de decisión de menor tamaño utilizando dichos subconjuntos.

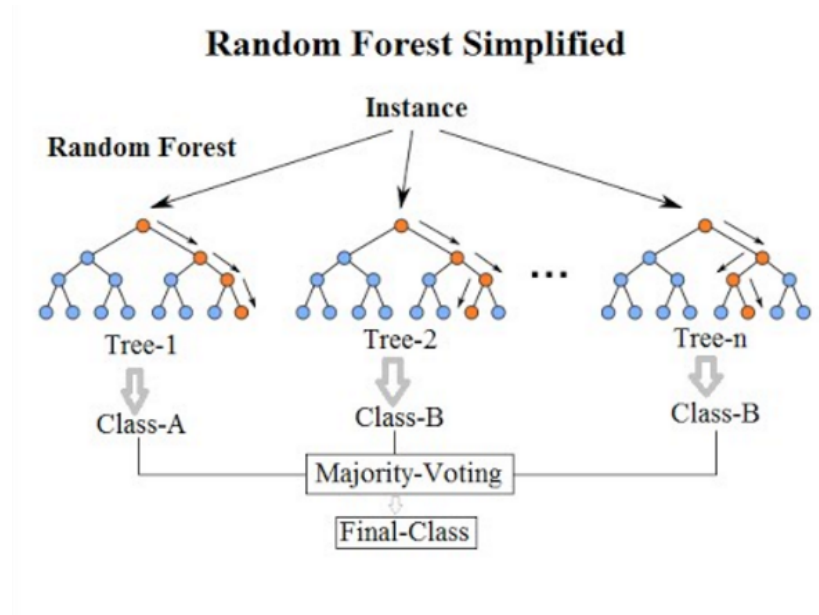


Figura 20. Random Forest Simplificado

Para entender el código :

La biblioteca *Sklearn*s utiliza una serie de hiperparámetros que se explican a continuación para poder entender el código del modelo Random Forest.

- **n_estimators**: indica el número de árboles de decisión que el algoritmo construye antes de tomar promedios de predicciones. Generalmente, un número elevado de árboles incrementa el rendimiento y conlleva a predicciones más estables. Sin embargo, también ralentiza el cálculo.
- **max_features**: es el número máximo de características que Random Forest tiene permitido intentar en un árbol individual.
- **min_sample_leaf**: determina el número mínimo de “hojas” que son necesarias para dividir un nodo interno.
- **n_jobs**: parámetro que indica al motor cuántos procesadores tiene permitido usar.
- **random_state**: hace que la salida del modelo sea replicable. Es decir, el modelo producirá siempre los mismos resultados cuando tenga un valor definido de `random_state` y siempre que se utilicen los mismos hiperparámetros y los mismos datos de entrenamiento.
- **oob_score**: se trata de un método de validación cruzada del Random Forest. En el muestreo, alrededor de un tercio de los datos no son utilizados para entrenar el modelo y por tanto pueden ser utilizados para evaluar su funcionamiento. Estas muestras se conocen como muestras “fuera de bolsa”.

Error

Respecto al error producido en el modelo Random Forest, se puede concluir que depende en gran medida de dos factores. Por un lado dicho error depende de la correlación existente entre cualquier par de árboles dentro del “bosque”. Aumentar la correlación aumenta la tasa de error. Por otro lado, la fuerza de cada árbol individual dentro del bosque afecta también a la tasa de error del modelo. Un árbol con una baja tasa de error es un clasificador fuerte. Aumentar la fuerza de un árbol individual disminuye la tasa de error del bosque.

Siendo M el número de variables de entrada del modelo, se especifica un número $m \ll M$ de forma que en cada nodo se seleccionen m variables de forma aleatoria dentro de M , y la mejor división de dichos m se utiliza para separar el nodo. El valor de m se mantiene constante a lo largo del crecimiento del bosque. Reduciendo m , se reduce al mismo tiempo la correlación y la fuerza. Aumentarlo implica aumentar ambas variables. En algún punto intermedio existe un rango “óptimo” de m , que suele ser generalmente bastante amplio. Es por ello que se utiliza `oob_score`, que obtiene un valor de m en un intervalo de manera rápida. Este es el único parámetro ajustable al cual el modelo es sensible.

Ventajas

El modelo de predicción Random Forest presenta numerosas ventajas frente a otros modelos predictivos:

- Es insuperable en precisión entre los algoritmos existentes actualmente.
- Se ejecuta eficientemente en grandes bases de datos.
- Permite manejar al mismo tiempo numerosas variables de entrada sin eliminación de las mismas.
- Estima qué variables son importantes para la clasificación.
- Genera una estimación interna imparcial del error de generalización a medida que avanza la construcción forestal.
- Cuenta con métodos para la estimación de los datos no existentes y mantiene la precisión a pesar de que falten gran parte de datos.
- Cuenta con métodos para equilibrar el error en la población de clases de conjuntos de datos no balanceados.
- Los bosques generados se pueden guardar para uso futuro en otros datos.
- Calcula prototipos que proporcionan información sobre la relación entre las variables y la clasificación.
- Calcula proximidades entre pares de casos que se pueden usar para agrupar, ubicar valores atípicos u ofrecer vistas interesantes de los datos.
- Las capacidades mencionadas pueden extenderse a los datos que no están etiquetados, lo cual lleva a un agrupamiento sin supervisión, vista de datos y detección de valores atípicos.
- Ofrece un método experimental para detectar interacciones variables.

Para mejorar un modelo de predicción machine Learning se recomienda:

- Usar más datos (de más calidad) y la ingeniería de características
- Afinar los hiperparámetros del algoritmo
- Probar con diferentes algoritmos.
- Una vez obtenidas las importancias de las variables, eliminar del modelo aquellas cuya importancia sea 0.

4.2.2.2 Modelo Prophet

El modelo Prophet se trata de un software de código abierto lanzado por Facebook. En particular fue creado por el equipo de Core de Data Science. Este modelo basado en un modelo aditivo sirve para predecir datos de series temporales. En dicho modelo las tendencias no lineales se ajustan a la estacionalidad anual, semanal y diaria, y es posible incorporar los efectos de las vacaciones y festivos. Además, dicho modelo es bastante robusto y eficiente para set de datos que contienen datos faltantes, cambios de tendencia y valores atípicos.

Para entender el modelo Prophet es necesario utilizar un modelo de series temporales que se puede descomponer en tres componentes principales del modelo: tendencia, estacionalidad y vacaciones.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

Donde:

- $g(t)$ representa la función de tendencia que modela los cambios no periódicos en el valor de las series temporales.
- $s(t)$ representa los cambios periódicos como por ejemplo la estacionalidad semanal o anual.
- $h(t)$ representa los efectos de las vacaciones que se producen en horarios potencialmente irregulares durante uno o más días.
- ϵ_t es el término del error, que representa cualquier cambio idiosincrásico que no sea acomodado por el modelo. Se asumirá que el parámetro del error será normalmente distribuido.

Utilizando el tiempo como regresor, Prophet ajusta varias funciones lineales y no lineales del tiempo como componentes. Con este método lo que se realiza es enmarcar el problema de predicción como un ejercicio de ajuste de curvas en lugar de considerar explícitamente la dependencia temporal de cada observación.

Para entender el código :

Algunos de los parámetros que se pueden modificar dentro del modelo Prophet son:

- Parámetros de tendencia:
 - **growth**: hace referencia al crecimiento, que puede seguir una tendencia “linear” o una tendencia “logística”.
 - **Changepoints**: lista de fechas en las cuáles se incluyen potenciales puntos de cambio de tendencia. Es un parámetro automático a menos que se especifique.
 - **n_changepoints**: Si los puntos de cambio no son especificados, se debe indicar el número de puntos de cambio para que se incluyan automáticamente.
 - **changepoint_prior_scale**: es un parámetro que sirve para cambiar la flexibilidad de la selección automática de puntos de cambio. Se utiliza para ajustar la intensidad de la dispersión de los datos, y evitar así que los cambios de tendencia estén sobre-ajustados (demasiada flexibilidad) o sub-ajustados (no tiene suficiente flexibilidad). Dicho parámetro por defecto tiene un valor de 0.05. Aumentarlo hace que la tendencia sea más flexible.
- Parámetros de estacionalidad y periodos vacacionales:
 - **yearly_seasonality**: ajusta la estacionalidad anual.
 - **weekly_seasonality**: ajusta la estacionalidad semanal.
 - **daily_seasonality**: ajusta la estacionalidad diaria.

- ***holidays***: alimenta un marco de datos que contiene el nombre y fecha de las vacaciones.
- ***seasonality_prior_scale***: cambia la fuerza e importancia del modelo de estacionalidad.
- ***holiday_prior_scale***: cambia la fuerza e importancia del modelo del periodo de vacaciones.

Ventajas:

Como se ha mencionado previamente, aplicando este modelo lo que se realiza es enmarcar el problema de pronóstico como un ejercicio de ajuste de curvas, que es intrínsecamente diferente de los modelos de serie temporales que cuentan explícitamente con la estructura de dependencia temporal de los datos. Las ventajas que esto presenta son:

- Se caracteriza por tener una gran flexibilidad. Es posible acomodar fácilmente la estacionalidad con múltiples periodos y permite al analista hacer distintas asunciones sobre las tendencias.
- Las medidas no necesitan estar regularmente espaciadas y no se necesita interpolar los valores faltantes. No es necesario eliminar los valores atípicos.
- La adaptación del modelo es muy rápida, lo que permite al analista explorar interactivamente muchas especificaciones de modelos.
- El modelo de predicción tiene parámetros que son fácilmente interpretables y que pueden ser modificados por el analista para imponer sus suposiciones sobre el pronóstico. Además es posible extender fácilmente el modelo para incluir nuevos componentes.
- Se trata de un modelo que realiza buenas predicciones para conjuntos de datos que contienen observaciones por hora, diarias y semanales con al menos unos pocos meses de histórico, pudiendo incluir fuertes estacionalidades múltiples de escala humana.
- El modelo detecta e incluye en su predicción días festivos y eventos importantes que ocurren a intervalos irregulares que se conocen de antemano, como por ejemplo, la Super Bowl.
- El modelo es capaz de reconocer tendencias que son curvas de crecimiento no lineales, donde una tendencia alcanza un límite natural o se satura.

4.2.3 Análisis de los resultados de los modelos de predicción

A continuación se muestran los resultados de precisión obtenidos para ambos modelos de predicción en el caso de los 6 restaurantes simulados. Estos resultados se han obtenido al realizar la predicción sobre el conjunto de datos de testeo. Es conveniente recordar que de los 13 meses de histórico que tenemos, se van a emplear 12 meses para entrenar los modelos y 1 mes para verificar la precisión de los mismos. En primer lugar se calcula la predicción de afluencia para las fechas del conjunto de datos de testeo y a continuación se compara con el valor real de afluencia para ese día que se muestra en el histórico. Esta comparación sirve para obtener la precisión del modelo.

Se observa que de media, el modelo Random Forest ofrece una precisión ligeramente superior. No obstante, esta diferencia es mínima. Además, ambos niveles de precisión son aceptables para interpretar los resultados como válidos y para considerarlos como predicción de afluencia de cada restaurante.

	Random Forest	Prophet
Restaurante 1	89,10%	90,34%
Restaurante 2	89,37%	89,93%
Restaurante 3	87,82%	87,39%
Restaurante 4	88,37%	89,11%
Restaurante 5	91,12%	84,82%
Restaurante 6	87,99%	89,17%
	88,96%	88,46%

Tabla 9. Resumen de precisión de los modelos para cada restaurante

La fórmula utilizada para calcular las precisiones mostradas del modelo para cada restaurante es la siguiente:

$$\%_{\text{precisión}} = \left(1 - \left| \frac{y_{\text{predicción}} - y_{\text{real}}}{y_{\text{real}}} \right| \right) \cdot 100$$

Esta fórmula también corresponde con:

$$\%_{\text{precisión}} = 100\% - MAPE$$

Siendo *MAPE* el error porcentual absoluto medio. Este valor mide el tamaño del error en términos de porcentaje, y se calcula de la siguiente forma:

$$MAPE = \left(\frac{1}{n} \sum \frac{|y_{\text{real}} - y_{\text{predicción}}|}{|y_{\text{real}}|} \right) \cdot 100$$

Una vez conocemos la precisión de cada modelo se procede a calcular la predicción de afluencia para el mes siguiente al último mes del histórico, que en este caso corresponde con el mes de Marzo.

La gráfica que muestra la predicción del conjunto de datos de testeo frente al valor real del mismo para el Restaurante 1 realizada con el modelo Random Forest es la siguiente:

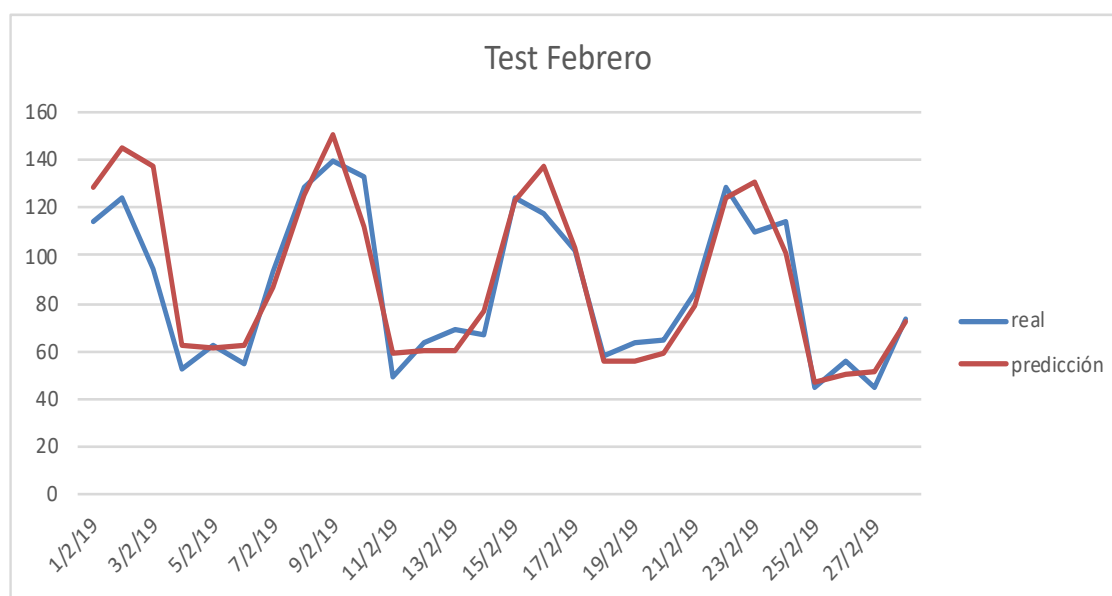


Figura 21. Test Random Forest Restaurante 1

Siendo el error de dicha predicción el que se muestra a continuación:

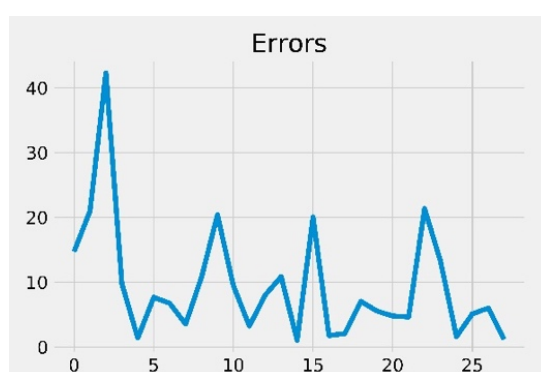
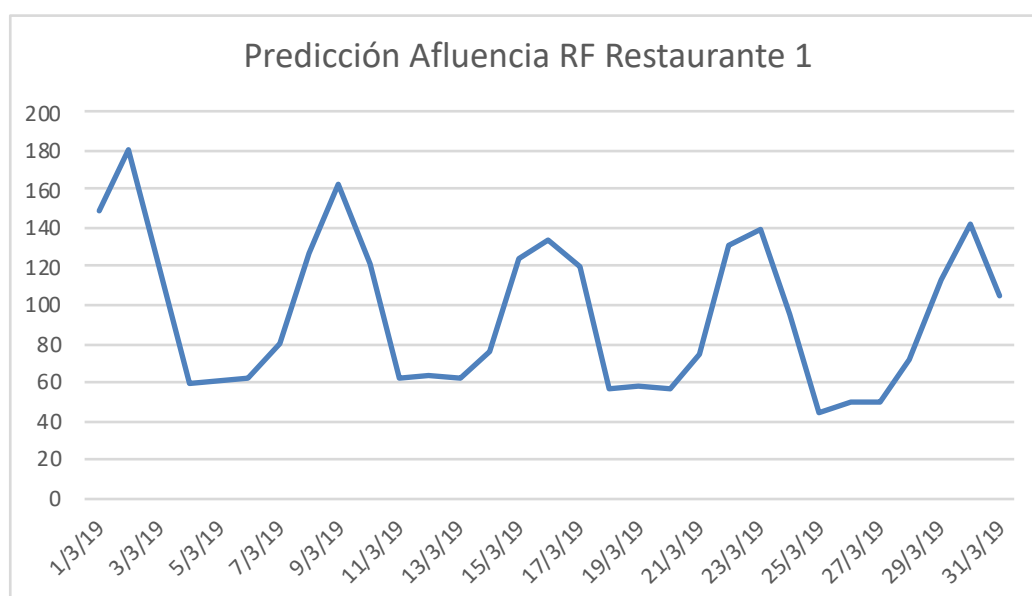


Figura 22. Error de predicción en el conjunto de datos de testeo Restaurante 1 (RF)

En este caso, la predicción para el mes de Marzo aportada por el modelo es:



Del mismo modo, el resultado obtenido para el conjunto de testeo del Restaurante 1 utilizando el modelo Prophet es el siguiente:

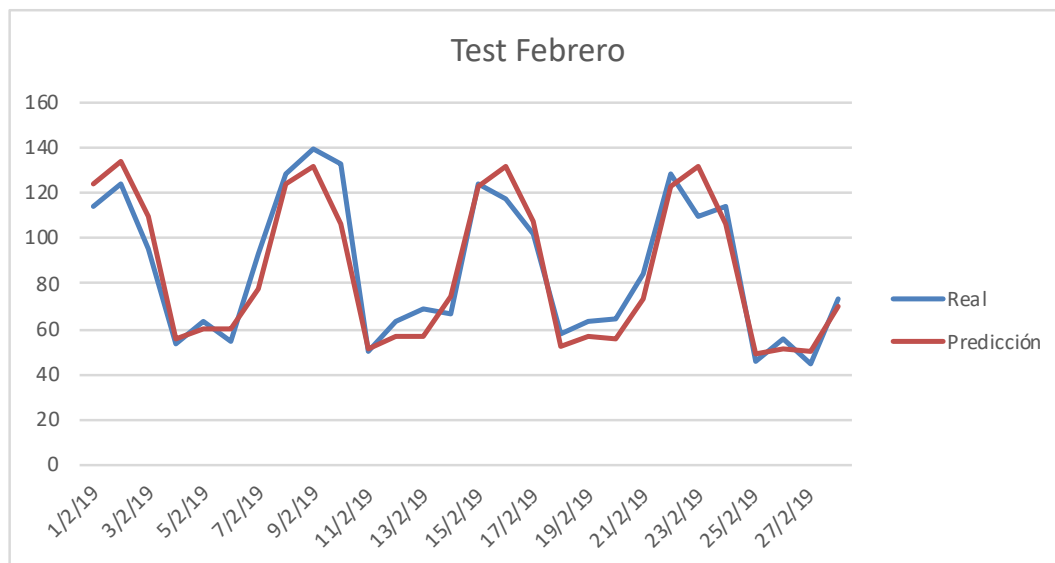


Figura 23. Test Prophet Restaurante 1

Por otro lado, la predicción del mes de Marzo obtenida para el mismo restaurante utilizando el modelo de predicción Prophet son:

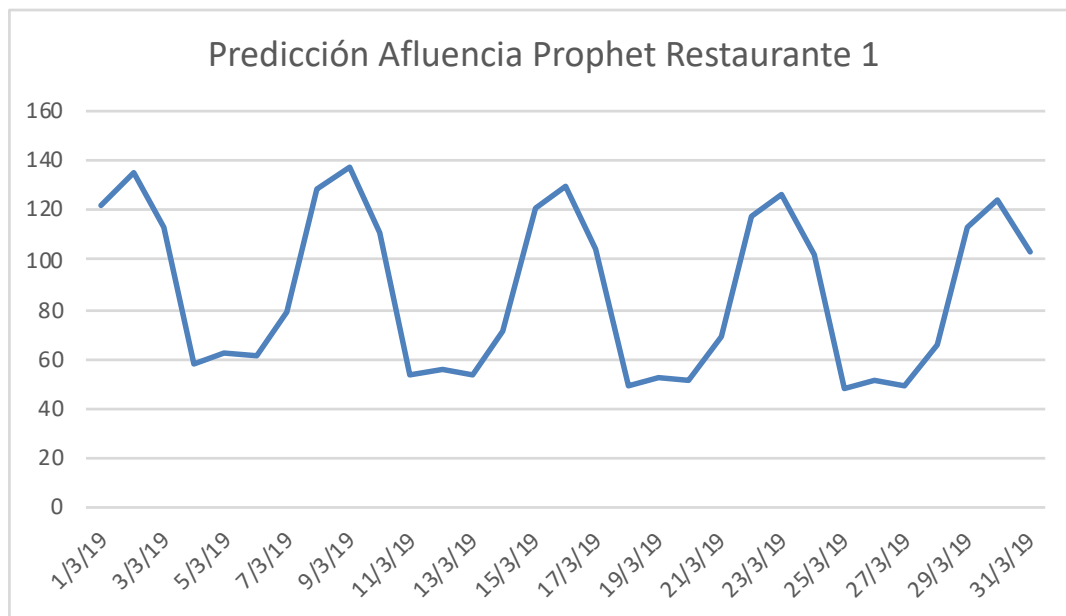


Figura 24. Predicción afluencia Restaurante 1 (Prophet)

Además, el modelo Prophet ofrece un análisis de los datos de entrenamiento donde se muestra la estacionalidad tanto diaria como anual y mensual, la tendencia y los periodos de vacaciones.

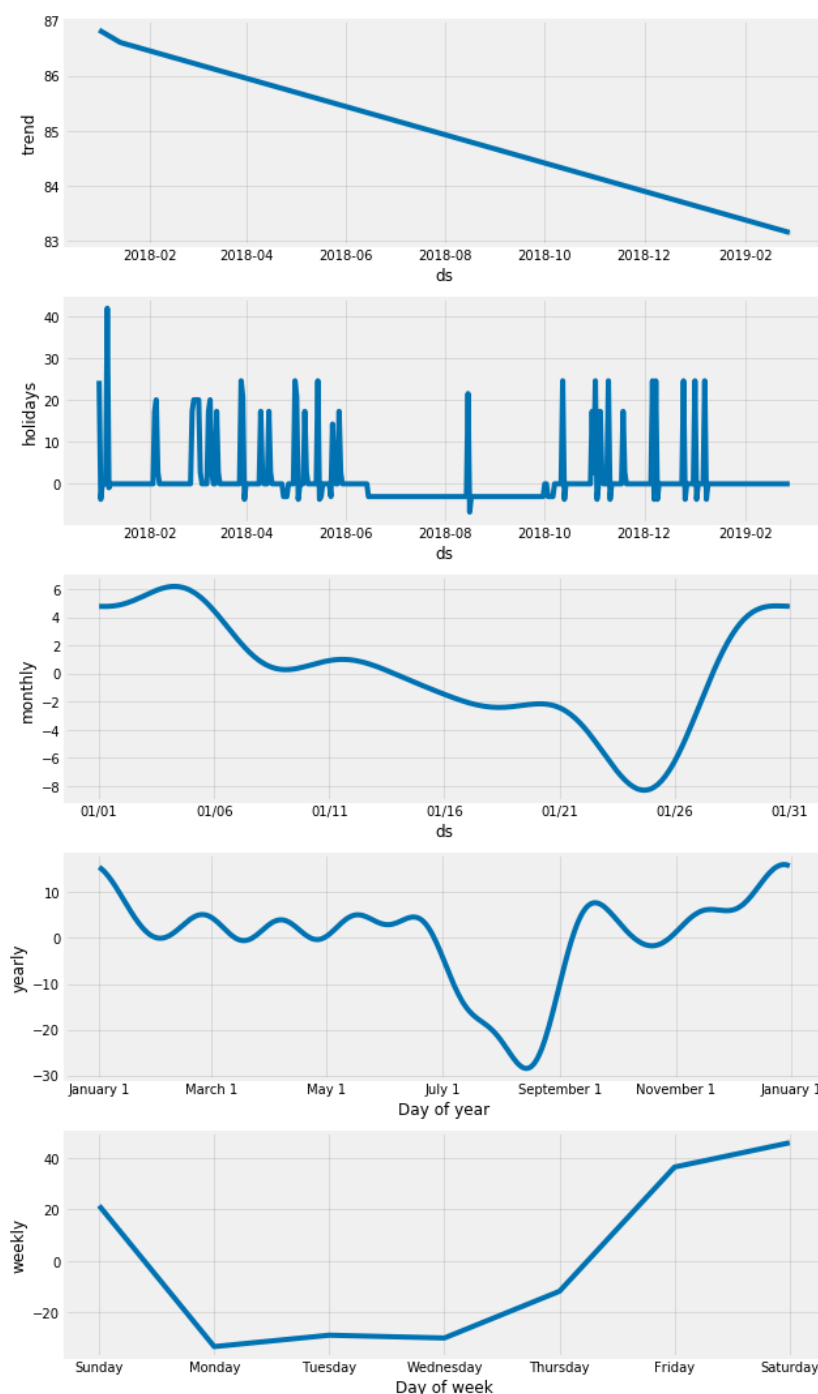


Figura 25. Propiedades de los datos (Prophet)

Como se puede observar en la imagen, hay una clara estacionalidad anual donde se produce un descenso de afluencia entre junio y septiembre. Además, se registran picos elevados de afluencia en el periodo de diciembre, octubre y enero. Del mismo modo, se registra mayor afluencia a principios de mes y esta va disminuyendo a lo largo del mismo. Semanalmente, los viernes, sábados y domingos son los días de la semana donde acude más gente al restaurante. Es preciso recordar para entender el comportamiento que este restaurante está situado en un centro comercial y es del tipo comida rápida. De modo que parece lógico pensar que durante el fin de semana tenga más afluencia que durante la semana. La afluencia durante los lunes, martes y miércoles parece mantenerse relativamente estable, comenzando a incrementar a partir del jueves.

Las gráficas obtenidas para los restaurantes 2,3,4,5 y 6 se han incluido en el Anexo D. Las gráficas correspondientes al conjunto de datos del testeo que reflejan la precisión de los modelos junto con la representación de la predicción del mes de Marzo se observan en el Anexo E.

4.2.4 Conclusiones

Observando el promedio global del porcentaje de precisión de ambos modelos destaca la ligera superioridad de precisión del modelo Random Forest (88,98 %) frente a la precisión del modelo Prophet (88,46 %).

Puesto que el conjunto de datos de testeo corresponde al mes de Marzo, mes en el cuál no hay eventos festivos ni vacaciones, los resultados entre ambos modelos son bastante similares. No obstante, para meses en los cuáles se producen dichos eventos es muy probable que el modelo Prophet ofrezca mejores resultados y precisión, ya que está especialmente preparado para detectar dichos eventos e incorporar su efecto en el modelo. Aunque no se han incluido los resultados, dicha funcionalidad ha sido verificada con los datos existentes.

No obstante, para el conjunto de validación la variación entre ambos modelos es muy ligera y podría utilizarse indistintamente uno u otro modelo, ofreciendo en ambos casos una buena precisión que permitirá predecir con éxito la afluencia que se producirá.

4.3 Optimización del número de empleados

4.3.1 Introducción

Una vez conocida la afluencia que cada local de la cadena de restauración va a tener, es posible utilizar esta información para calcular el número de empleados más adecuado que se va a necesitar en cada restaurante y en cada momento para ofrecer un servicio óptimo a los clientes y para mantener una buena calidad de trabajo de los empleados.

Para realizar la optimización se han realizado diversas hipótesis y se ha aplicado un modelo matemático de optimización implementado en Python. Para ello se ha utilizado un optimizador conocido como *Gurobi*. Este solucionador permite resolver todo tipo de problemas de optimización: problemas lineales (LP), problemas lineales enteros mixtos (MILP), problemas cuadráticos enteros mixtos (MIQP), problemas cuadráticos (QP), problemas restringidos cuadráticamente (QCP) y problemas enteros mixtos restringidos cuadráticamente (MIQCP).

4.3.2 Desarrollo del modelo de optimización

Para empezar, para poder calcular el número óptimo de empleados por cada hora es necesario previamente calcular el número de comensales que acudirá al restaurante en cada franja horaria. El resultado obtenido de la predicción de afluencia es el número total de comensales que acudirán al restaurante en un día determinado, de modo que es preciso discretizar este dato y obtener así la repartición de los comensales a lo largo del día por franjas horarias.

Para realizarlo, se acude a la información proporcionada por las analíticas de google sobre la distribución de afluencia en restaurantes que representan a aquellos que se han simulado en este proyecto. Dichos factores variarán en función de la localización de los restaurantes. En este caso se ha asumido para poder obtener los factores de afluencia que los locales de centro comerciales (Restaurante 1 y 4) se encontrarán en *La Vaguada, Madrid*, los restaurantes de barrios residenciales (Restaurante 2 y 5) estarán en *Goya, Madrid*, y los restaurantes de zonas de oficinas se situarán en la *Zona de las Cuatro Torres en Castellana 280, Madrid*.

De modo que, los factores de afluencia que servirán para discretizar el valor de afluencia diaria de cada restaurante son los siguientes:

Restaurante 1	CC	L	M	X	J	V	S	D
13:00	13:30	0,08	0,06	0,07	0,06	0,08	0,06	0,07
13:30	14:00	0,09	0,08	0,1	0,08	0,08	0,08	0,13
14:00	14:30	0,1	0,13	0,13	0,1	0,1	0,14	0,17
14:30	15:00	0,17	0,15	0,16	0,12	0,12	0,17	0,15
15:00	15:30	0,18	0,17	0,16	0,14	0,13	0,17	0,15
15:30	16:00	0,15	0,17	0,15	0,18	0,14	0,16	0,13
16:00	16:30	0,13	0,13	0,13	0,17	0,17	0,12	0,1
16:30	17:00	0,1	0,11	0,1	0,15	0,18	0,1	0,1

Figura 26. Factores de afluencia para el Restaurante 1 (localizado en centro comercial)

Restaurante 2	R	L	M	X	J	V	S	D
13:00	13:30	0,1	0,1	0,1	0,08	0,07	0,09	0,06
13:30	14:00	0,15	0,17	0,12	0,15	0,12	0,12	0,1
14:00	14:30	0,18	0,23	0,14	0,22	0,25	0,18	0,16
14:30	15:00	0,25	0,23	0,2	0,25	0,25	0,18	0,24
15:00	15:30	0,2	0,17	0,22	0,2	0,17	0,25	0,26
15:30	16:00	0,12	0,1	0,22	0,1	0,14	0,18	0,18

Figura 27. Factores de afluencia para el Restaurante 2 (localizado en barrio residencial)

Restaurante 3	O	L	M	X	J	V	S	D
12:00	12:30	0,07	0,04	0,05	0,02	0,05	0,1	0,06
12:30	13:00	0,08	0,07	0,05	0,06	0,06	0,1	0,07
13:00	13:30	0,1	0,12	0,12	0,08	0,07	0,11	0,07
13:30	14:00	0,15	0,17	0,16	0,12	0,08	0,12	0,08
14:00	14:30	0,2	0,2	0,18	0,16	0,14	0,15	0,15
14:30	15:00	0,2	0,17	0,18	0,21	0,2	0,15	0,2
15:00	15:30	0,1	0,14	0,16	0,2	0,21	0,15	0,2
15:30	16:00	0,1	0,09	0,1	0,15	0,19	0,12	0,17

Figura 28. Factores de afluencia para el Restaurante 3 (localizado en zona de oficina)

Restaurante 4	CC	L	M	X	J	V	S	D
13:00	13:30	0,08	0,06	0,07	0,06	0,08	0,06	0,07
13:30	14:00	0,09	0,08	0,1	0,08	0,08	0,08	0,13
14:00	14:30	0,1	0,13	0,13	0,1	0,1	0,14	0,17
14:30	15:00	0,17	0,15	0,16	0,12	0,12	0,17	0,15
15:00	15:30	0,18	0,17	0,16	0,14	0,13	0,17	0,15
15:30	16:00	0,15	0,17	0,15	0,18	0,14	0,16	0,13
16:00	16:30	0,13	0,13	0,13	0,17	0,17	0,12	0,1
16:30	17:00	0,1	0,11	0,1	0,15	0,18	0,1	0,1

Figura 29. Factores de afluencia para el Restaurante 4 (localizado en centro comercial)

Restaurante 5	R	L	M	X	J	V	S	D
13:00	13:30	0,05	0,04	0,03	0,05	0,07	0,04	0,04
13:30	14:00	0,09	0,05	0,05	0,11	0,08	0,09	0,06
14:00	14:30	0,16	0,18	0,11	0,2	0,17	0,17	0,13
14:30	15:00	0,18	0,23	0,17	0,25	0,22	0,17	0,23
15:00	15:30	0,22	0,23	0,21	0,2	0,22	0,24	0,28
15:30	16:00	0,16	0,13	0,21	0,09	0,13	0,18	0,18
16:00	16:30	0,09	0,1	0,17	0,06	0,06	0,08	0,06
16:30	17:00	0,05	0,04	0,05	0,04	0,05	0,03	0,02

Figura 30. Factores de afluencia para el Restaurante 5 (localizado en barrio residencial)

Restaurante 6	O	L	M	X	J	V	S	D
12:00	12:30	0,07	0,04	0,05	0,02	0,05	0,1	0,06
12:30	13:00	0,08	0,07	0,05	0,06	0,06	0,1	0,07
13:00	13:30	0,1	0,12	0,12	0,08	0,07	0,11	0,07
13:30	14:00	0,15	0,17	0,16	0,12	0,08	0,12	0,08
14:00	14:30	0,2	0,2	0,18	0,16	0,14	0,15	0,15
14:30	15:00	0,2	0,17	0,18	0,21	0,2	0,15	0,2
15:00	15:30	0,1	0,14	0,16	0,2	0,21	0,15	0,2
15:30	16:00	0,1	0,09	0,1	0,15	0,19	0,12	0,17

Figura 31. Factores de afluencia para el Restaurante 6 (localizado en zona de oficinas)

Para obtener un ejemplo de una mejor visualización de la distribución los factores se muestran las siguientes gráficas correspondientes al Restaurante 1 y 4, situados en un Centro Comercial:

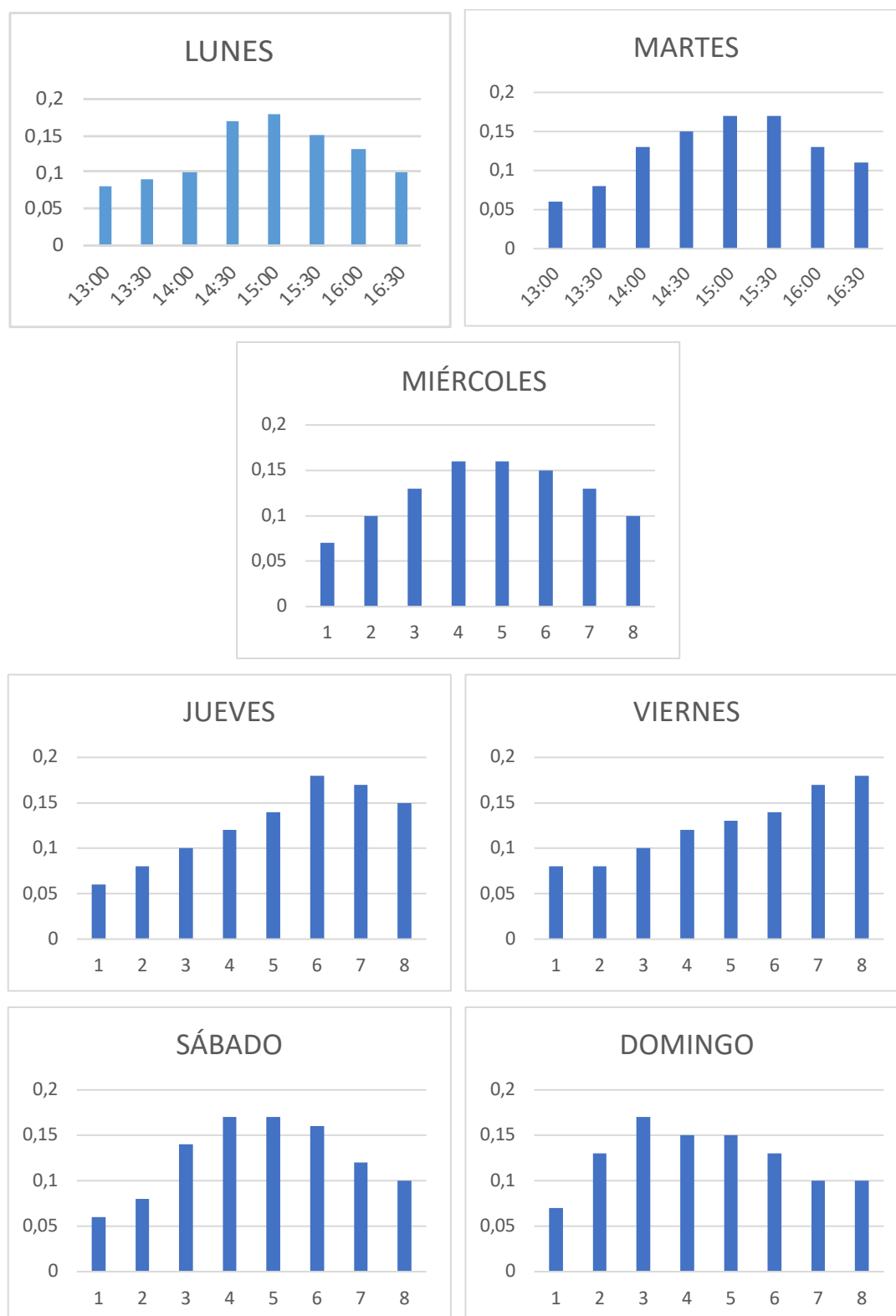


Figura 32. Visualización gráfica de factores de afluencia para el Restaurante 1

Para visualizar el resto de gráficas de distribución de los factores de afluencia ver Anexo F.

Tras aplicar los factores de afluencia correspondientes a cada día de la predicción, obtendremos la repartición diaria por restaurantes del número de comensales. Una vez conocido este valor se procede a comenzar con la optimización. El modelo matemático se desarrollará en Python como se ha mencionado previamente.

El siguiente paso es establecer las variables que se van a calcular en la optimización y las restricciones que van a ser necesarias para realizar el modelo matemático de optimización.

De este modelo se pretende obtener el número de empleados óptimo para satisfacer la demanda pero también la demanda perdida que se producirá en el local con el número de empleados que se obtiene. Es importante encontrar el equilibrio entre el coste de añadir un nuevo empleado en el local y el coste de perder un número de comensales. A continuación se muestra la porción de código donde se define el modelo matemático junto con las variables que se van a calcular en la optimización.

```
1. #Name of model
2. m = Model('Opt_empleados')
3.
4. #variables
5. empleados = m.addVars(dia, hora, vtype=GRB.INTEGER, name="Empleados")
6. demanda_perdida = m.addVars(dia, hora, vtype = GRB.INTEGER, name = "DemandaPerdida")
   #numero de comensales perdidos
7.
```

Ambas variables dependerán del día y de la hora y serán del tipo entero.

Por otro lado, se han definido los parámetros que formarán parte del modelo:

```
1. #Parameters
2. precio_empleados = 11
3. margen_venta = 9.375
4. ratio_comensales_max = 15
5. precio_perdida = 18.75
```

Para calcular dichos parámetros se han realizado una serie de cálculos e hipótesis. En primer lugar, el coste de los empleados se ha estimado teniendo en cuenta el salario mensual, el plus de transporte mensual, la nocturnidad, las horas extras, los procesos de formación y la vestimenta y material necesarios.

Teniendo en cuenta que el salario mensual de los empleados de los restaurantes simulados corresponde con el salario medio de un camarero trabajando en un restaurante de una cadena de restauración como por ejemplo la cadena Grupo Vips, el salario mensual es de 996,8 €/mes. El plus transporte sería de 80 € mensuales. Las horas extras se pagan un 75% más caras que el precio de la hora correspondiente al salario asignado, mientras que las horas de nocturnidad se pagan un 25% más caras. Además, la formación, vestimenta y material incrementa el coste final del empleado por hora en un 10%. Se ha estimado que los empleados trabajaran 40 horas semanales, pudiendo realizar un máximo de 15 horas extras al mes. Teniendo en cuenta toda esta información se obtiene que el coste total de un empleado por hora es de 11 €/hora.

Por otro lado, para calcular el margen de venta de los productos que se venderán en el local se ha tenido en cuenta el hecho de que para fijar el precio final de los platos que serán vendidos en el restaurante se incrementa el coste total del producto en un 60%. De esta forma se ha realizado una estimación de cuál será el coste de elaboración de un producto medio para un restaurante tipo dentro de una cadena de restauración y se han dividido los costes en dos tipos: variables y fijos. Dentro de los costes variables se incluye el coste de la materia prima. Estos costes supondrán un 40% del coste total. Dentro de los costes fijos se incluye el alquiler del local, la energía eléctrica, el agua, el personal (camareros, cocineros, limpieza, central...), tributos y el material (uniforme, utensilios...), costes que supondrán el 60% de los costes totales. Tras un análisis exhaustivo sobre este tipo de negocios se ha observado que el precio medio de venta por servicio en este tipo de restaurantes es de 25 €, lo que significa que de este precio, 15,63 € corresponden al coste del producto y 9,37€ corresponden al margen de venta.

Respecto al coste de perder un comensal, teniendo en cuenta el dato presentado al principio del proyecto que afirma que el 96% ²de los clientes no satisfechos no pondrán una queja, pero el 91% de los mismos se marcharán y no volverán a dicho restaurante, se ha asumido que la pérdida de un comensal corresponde con la pérdida del margen que habrían ingresado en esa visita (9,37€) más la pérdida del margen que habrían ingresado en la siguiente puesto que no se va a producir (9,37€). En total esto supondría un valor de 18,75€. Cabe destacar que el 87% de la demanda perdida corresponde con el no retorno de clientes insatisfechos, mientras que el 13% corresponde con la disminución de la frecuencia de los mismos. Como se puede observar, resulta mucho más perjudicial para el restaurante la pérdida de un cliente que el coste de añadir un empleado, según las hipótesis planteadas. Además, el coste de perder un cliente incluye un coste intangible pero muy importante a considerar puesto que afectará a la imagen de marca y puesto que un cliente no satisfecho contará su experiencia a entre 9 y 15 personas. Alrededor del 13% de los clientes no satisfechos se lo contarán a más de 20 personas. De modo que esto puede tener un gran impacto sobre el negocio y la reputación del mismo, que se puede ver traducido en importantes pérdidas. Es por ello que el valor de pérdida de un comensal es tan elevado y es de vital importancia evitarlo.

El último parámetro que queda por determinar es el ratio de comensales por camarero. El objetivo del programa es que este parámetro se pueda modificar en función de la exigencia del local y del nivel de atención al cliente y calidad de servicio que se requiera en cada caso. Tras realizar un análisis de sensibilidad y teniendo en cuenta las características de los locales simulados y el servicio requerido en dichos locales se ha estimado que el valor óptimo para este caso es de 15 comensales por camarero por hora. De esta forma se asegura un equilibrio entre las condiciones de trabajo de los empleados y una buena calidad de servicio, evitando así perder demanda. No obstante, el programa ofrece la posibilidad de ajustar este parámetro según las necesidades y ejecutar el programa para poder obtener el nuevo número de empleados óptimo si se desea.

El programa ofrece también la posibilidad de fijar un máximo de empleados, de forma que se puedan ajustar los horarios y turnos a los empleados disponibles. No obstante se ha fijado en este caso un número máximo de empleados muy elevado para este tipo de locales (50 empleados) con el objetivo de calcular el número exacto de personal que se necesitaría para no perder demanda y ofrecer un servicio óptimo.

El siguiente paso consiste en fijar las restricciones del modelo que servirán para ejecutar la optimización de manera correcta. En este caso se han fijado cuatro restricciones, pero es posible readaptar este modelo una vez se empiecen las negociaciones con un cliente real para poder incorporar nuevas restricciones propias de cada local en particular. De ahí la flexibilidad de esta herramienta, ya que permitirá una completa adaptación del modelo a cada local y sus características fácilmente.

En este caso la primera restricción pretende que se calcule un número de empleados por día y por hora que sea siempre inferior al número máximo de empleados disponibles en la bolsa de trabajo de la cadena de restauración. La segunda restricción sirve para asegurar que el número de empleados que se calcule teniendo en cuenta el ratio de comensales que puede atender sea suficiente para satisfacer la demanda, pudiendo existir la posibilidad de que se pierdan comensales por exceso de afluencia. Esta restricción permitirá calcular hasta qué punto es conveniente incrementar un empleado frente a perder a un cliente. La tercera restricción hace posible que haya demanda perdida dentro del modelo, definiendo que esta puede ser mayor o igual que cero. La última restricción aplica el mismo procedimiento a los empleados, evitando

² Source: 1st Financial Training services

que el modelo obtenga un número de empleados negativo. Es preciso definir estas últimas restricciones ya que el modelo de optimización no es capaz de identificar en qué consiste cada variable en el mundo real y los resultados podrían salir negativos si con ello se optimizara el resultado.

```
1. #Add constraints
2. m.addConstrs((empleados[d,h] <= maximo[h] for d in dia for h in hora),name="constremp
   leado")
3. m.addConstrs((empleados[d,h]*ratio_comensales_max >= (comensales[d]*curvas[d,h] - dem
   anda_perdida[d,h]) for d in dia for h in hora ),name = "-")
4. m.addConstrs((demanda_perdida[d,h] >= 0 for d in dia for h in hora), name = "constrde
   manda")
5. m.addConstrs((empleados[d,h] >= 0 for d in dia for h in hora), name = "constrempleado
   2")
6.
```

La última parte del proceso consiste en fijar el objetivo de la optimización. En este caso el objetivo es maximizar los beneficios del restaurante habiendo asegurado que la calidad de servicio es la óptima y deseada en el ratio de comensales por empleado y tras haber fijado las restricciones con las que será posible atender el mayor número de comensales asegurando un equilibrio entre empleados y demanda perdida para obtener el mayor beneficio. La función beneficio consiste en calcular los ingresos que supondrían los comensales que acudirán al restaurante menos los gastos y costes tanto de empleados como de pérdida de comensales.

```
1. #the objective is to maximize profit
2. m.setObjective(quicksum((comensales[d]*precio_venta - precio_empleados*empleados[d,h]
   - demanda_perdida[d,h]*precio_perdida) for d in dia for h in hora), GRB.MAXIMIZE)
```

Una vez definida la función objetivo del problema de optimización que en este caso es una maximización se procede a ejecutar la optimización, que tendrá en cuenta todos los aspectos mencionados previamente y permitirá obtener los resultados óptimos de las dos variables definidas: el número de empleados para cada local y la demanda perdida de los mismos.

```
1. #Optimization
2. m.optimize()
```

4.3.3 Análisis de los resultados

Puesto que en el proyecto se han utilizado dos modelos matemáticos de predicción de afluencia, el modelo de optimización se ha aplicado dos veces, una para la predicción aportada por el modelo Random Forest y otra para la predicción aportada por el modelo Prophet. A continuación se mostrarán algunos ejemplos de los resultados obtenidos. El formato de resultados será el mismo para ambos casos.

Cabe destacar, que a pesar de que las franjas horarias obtenidas a partir de los factores de afluencia y utilizadas para la optimización han sido de media hora, en los resultados para observar el número de empleados se han agrupado las franjas para establecer turnos de una hora en los restaurantes. De esta forma, si en la optimización se obtenían ocho franjas de media hora, en la visualización de los resultados se mostrarán tan solo cuatro franjas de una hora, obtenidas tras haber adicionado los resultados de las franjas de media hora consecutivas.

A continuación se muestran los resultados obtenidos para el Restaurante 1, con los resultados de predicción obtenidos del modelo Random Forest:

día	hora	empleados caso base	empleados óptimo	visitas comensales	capacidad de atender comensales	diferencia	ratio comensales	ratio estandar	productividad 1 (clientes por hora-persona trabajada)	productividad 2 (ingresos por cada coste)
d1	h1	2	2	24	30	6	12	15	12	27
d1	h2	2	3	33	45	12	11	15	11	25
d1	h3	2	3	41	45	4	14	15	14	31
d1	h4	2	4	53	60	7	13	15	13	30
d2	h1	2	2	26	30	4	13	15	13	30
d2	h2	2	4	56	60	4	14	15	14	32
d2	h3	2	4	60	60	0	15	15	15	34
d2	h4	2	3	40	45	5	13	15	13	30
d3	h1	2	2	25	30	5	13	15	13	28
d3	h2	2	3	40	45	5	13	15	13	30
d3	h3	2	3	35	45	10	12	15	12	27
d3	h4	2	2	25	30	5	13	15	13	28
d4	h1	2	1	11	15	4	11	15	11	25
d4	h2	2	2	17	30	13	9	15	9	19
d4	h3	2	2	20	30	10	10	15	10	23
d4	h4	2	1	14	15	1	14	15	14	32
d5	h1	2	1	9	15	6	9	15	9	20
d5	h2	2	2	18	30	12	9	15	9	20
d5	h3	2	2	21	30	9	11	15	11	24
d5	h4	2	1	15	15	0	15	15	15	34
d6	h1	2	1	11	15	4	11	15	11	25
d6	h2	2	2	18	30	12	9	15	9	20
d6	h3	2	2	20	30	10	10	15	10	23
d6	h4	2	1	15	15	0	15	15	15	34
d7	h1	2	1	12	15	3	12	15	12	27
d7	h2	2	2	18	30	12	9	15	9	20
d7	h3	2	2	26	30	4	13	15	13	30
d7	h4	2	2	26	30	4	13	15	13	30

Figura 33. Resultado optimización Restaurante 1

En la tabla se muestran los resultados divididos por día y hora. En este caso, el *d1* corresponde con el primer día del mes de predicción, que es el 1 de Marzo de 2019. Se ha mostrado tan solo la primera semana para que sirva como ejemplo, pero la tabla continuaría hasta el día 31 de predicción.

En este caso, el Restaurante 1 posee una capacidad de 150 comensales por día. Se ha asumido como caso base que en restaurantes con capacidad de 150 comensales trabajarán 2 empleados, mientras que en aquellos con capacidad 80 trabajará 1 solo empleado. La tabla mostrada anteriormente servirá como ejemplo de optimización para los Restaurantes 2 y 3 puesto que los Restaurantes 2 y 3 tienen la misma capacidad que el Restaurante 1. A continuación se va a mostrar un ejemplo de optimización del Restaurante 4, cuya capacidad es de 80 comensales. Los Restaurantes 5 y 6 tendrán también capacidad de 80 comensales y por tanto un empleado de servicio.

En este caso, para la optimización de empleados para el Restaurante 6 mostrada a continuación se han empleado los datos procedentes del modelo Prophet, como ejemplo.

día	hora	empleados caso base	empleados	visitas comensales	capacidad de atender comensales	diferencia	ratio comensales	ratio estandar	productividad 1 (clientes por hora-persona trabajada)	productividad 2 (ingresos por cada coste)
d1	h1	1	1	8	15	7	8	15	8	18
d1	h2	1	1	10	15	5	10	15	10	23
d1	h3	1	1	14	15	1	14	15	14	32
d1	h4	1	2	20	30	10	10	15	10	23
d2	h1	1	1	4	15	11	4	15	4	9
d2	h2	1	1	4	15	11	4	15	4	9
d2	h3	1	1	8	15	7	8	15	8	18
d2	h4	1	1	10	15	5	10	15	10	23
d3	h1	1	1	3	15	12	3	15	3	7
d3	h2	1	1	4	15	11	4	15	4	9
d3	h3	1	1	8	15	7	8	15	8	18
d3	h4	1	1	11	15	4	11	15	11	25
d4	h1	1	1	7	15	8	7	15	7	16
d4	h2	1	1	14	15	1	14	15	14	32
d4	h3	1	2	28	30	2	14	15	14	32
d4	h4	1	2	21	30	9	10,5	15	10,5	24
d5	h1	1	1	7	15	8	7	15	7	16
d5	h2	1	2	17	30	13	8,5	15	8,5	19
d5	h3	1	2	27	30	3	13,5	15	13,5	31
d5	h4	1	2	17	30	13	8,5	15	8,5	19
d6	h1	1	1	12	15	3	12	15	12	27
d6	h2	1	2	18	30	12	9	15	9	20
d6	h3	1	2	18	30	12	9	15	9	20
d6	h4	1	1	12	15	3	12	15	12	27
d7	h1	1	1	10	15	5	10	15	10	23
d7	h2	1	1	15	15	0	15	15	15	34
d7	h3	1	2	23	30	7	11,5	15	11,5	26
d7	h4	1	2	18	30	12	9	15	9	20

Figura 34. Resultados optimización Restaurante 6

Es preciso aclarar, que la optimización se ha ejecutado para los seis restaurantes existentes y se ha repetido el proceso para cada restaurante utilizando como datos de entrada por un lado la predicción del modelo Prophet y la predicción del modelo Random Forest. El objetivo es obtener todos los resultados para todas las posibilidades existentes y poder incorporar esta información en la herramienta donde se visualizará la información de forma que el usuario pueda decidir qué datos de entrada prefiere utilizar y que se muestren los resultados correspondientes. En el siguiente apartado se analizará más en profundidad las consecuencias de aplicar en el negocio los resultados obtenidos de esta optimización.

No obstante, se puede observar en ambos casos que existe variación entre el número de empleados que están trabajando actualmente, mostrados en la columna de *Caso Base* y el número de empleados óptimo calculado. En la tabla de resultados también se muestra cuál es la capacidad de atender a comensales en el caso óptimo, y cuál es la diferencia entre la capacidad de atender comensales y los comensales que realmente atenderán según la predicción. Esto permitirá saber con qué margen contamos en caso de que acudan más o menos comensales de los esperados. Por otro lado se muestra el ratio de comensales actual si acudieran el número previsto de comensales respecto al número de empleados que se han calculado como necesarios, y el ratio de comensales base que se ha empleado para la optimización. Ya se ha mencionado que este valor podría modificarse. Por último, se muestran dos medidas de productividad calculadas: la primera corresponde a la cantidad de comensales que atenderá cada empleado por hora, y la segunda la cantidad de ingresos frente a los gastos con ese número de empleados. Para calcular este último valor se han empleado los valores de los parámetros definidos en el modelo.

Cabe destacar que en el modelo ejecutado se ha obtenido demanda perdida nula en todos los casos. Es por ello que no se refleja en la tabla de resultados. El modelo ha concluido que será menos costoso para la empresa contratar nuevos empleados que puedan ayudar a satisfacer la demanda frente a tener demanda perdida. Además, el hecho de tener empleados cuyo ratio de comensales sea menor que el establecido permitirá en ciertos casos ofrecer mayor calidad aún,

y capacidad de ofrecer servicios extras o de personalización de oferta. Al mismo tiempo, al conocer la capacidad de ocupación de cada empleado, se podrán reasignar determinadas tareas a aquellos que estén menos ocupados, optimizando así la gestión completa del restaurante.

No obstante, los resultados mostrados supondrán una recomendación para los empresarios. En el siguiente apartado se mostrará el impacto en los beneficios y en la gestión del personal de aplicar dichos resultados a los distintos locales de la cadena de restauración.

4.4 Herramienta

En este apartado se presenta la herramienta final que se ha elaborado a lo largo del proyecto. Dicha herramienta está compuesta por distintas interfaces que corresponden con las tres partes en las que está estructurado el proyecto: Histórico, Predicción de afluencia y Optimización de empleados.

En la primera interfaz, además del histórico de cada restaurante es posible observar ciertos *KPIs* (Key performance indicators) o indicadores claves de rendimiento con los que podremos tener una visión de cómo los restaurantes están funcionando y detectar así fácilmente dónde se debe focalizar la atención para mejorar el rendimiento. Además, cada interfaz ofrece la posibilidad de seleccionar ciertos filtros de forma que la información que se muestra se pueda personalizar en tiempo real y según las necesidades. De esta manera, una cadena de restaurantes podrá tener resumida la información global del conjunto de todos sus restaurantes, pero podrá también seleccionar si lo desea un restaurante en particular y la información mostrada se actualizará con aquella que corresponda a los filtros seleccionados.

La segunda interfaz corresponde con la información de la predicción de comensales que va a haber en cada local. De la misma forma, existen una serie de filtros que se pueden seleccionar para personalizar la visualización. Además, en dicha interfaz será posible observar los resultados procedentes de ambos modelos matemáticos de predicción utilizados. La idea es que el propio usuario pueda elegir qué modelo le conviene más en función de la precisión y circunstancia para cada ocasión.

En la tercera y última interfaz se muestran los resultados de la optimización de empleados. En esta pantalla será posible visualizar la predicción calculada de distintas formas: como promedio de ambos modelos o importando directamente los resultados de uno de los modelos si se desea. A continuación se mostrará una comparativa de dos casos. El primer caso corresponde al caso base, donde se mostrará el número de empleados establecido de forma habitual: 2 empleados para locales con capacidad de 150 comensales, 1 empleado para locales con capacidad de 80 comensales. El segundo caso se trata del caso óptimo, donde el programa calcula el número de empleados necesarios para satisfacer la demanda que va a haber en el restaurante, manteniendo un nivel de calidad de servicio y asegurando buenas condiciones de trabajo para el empleado. Se mostrarán los resultados económicos de utilizar los empleados mencionados en ambos casos de forma que se pueda comparar fácilmente el beneficio que se obtendría en cada caso, y poder así tomar decisiones en consecuencia.

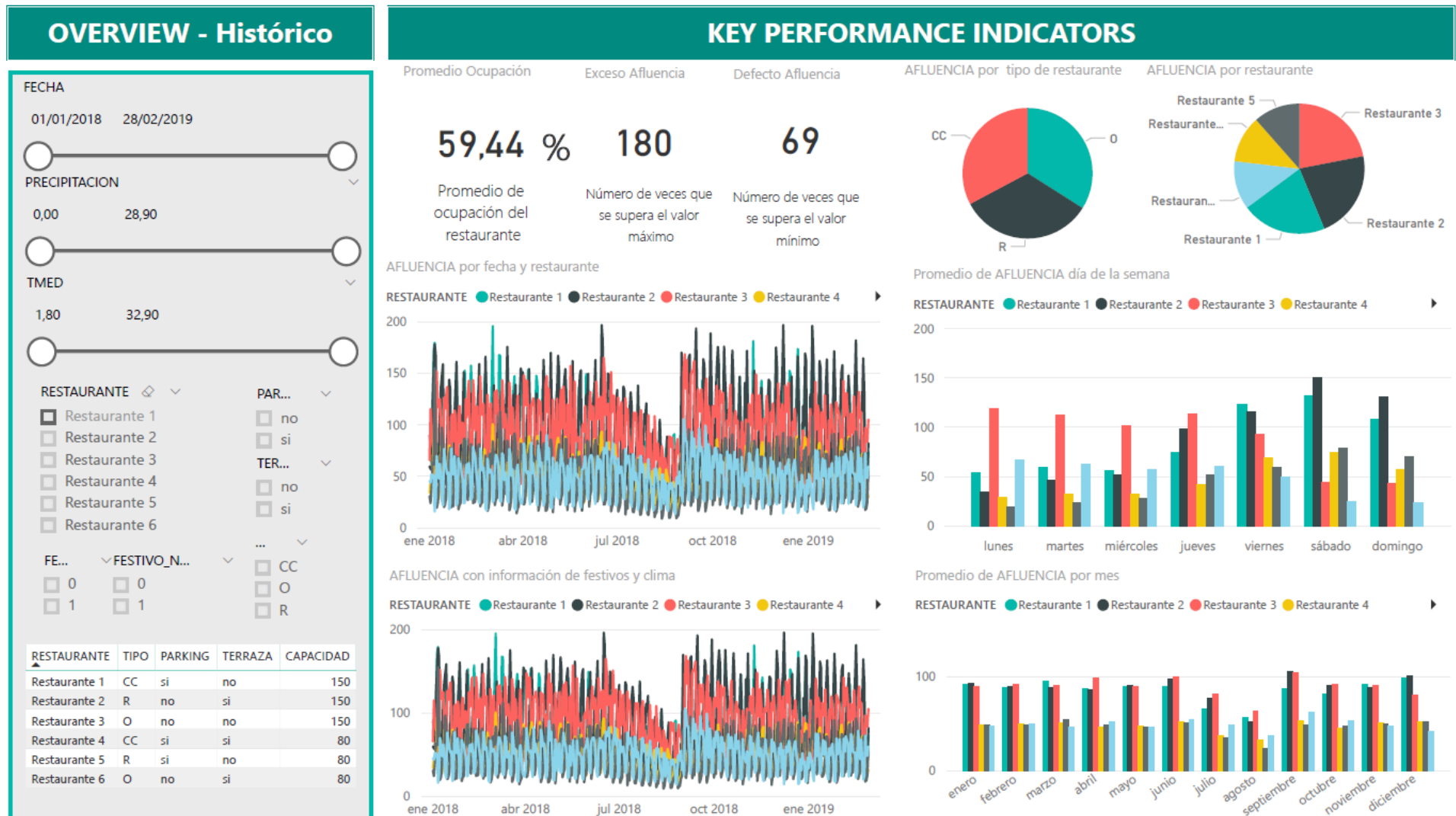


Figura 35. Primera interfaz de la herramienta (Histórico)

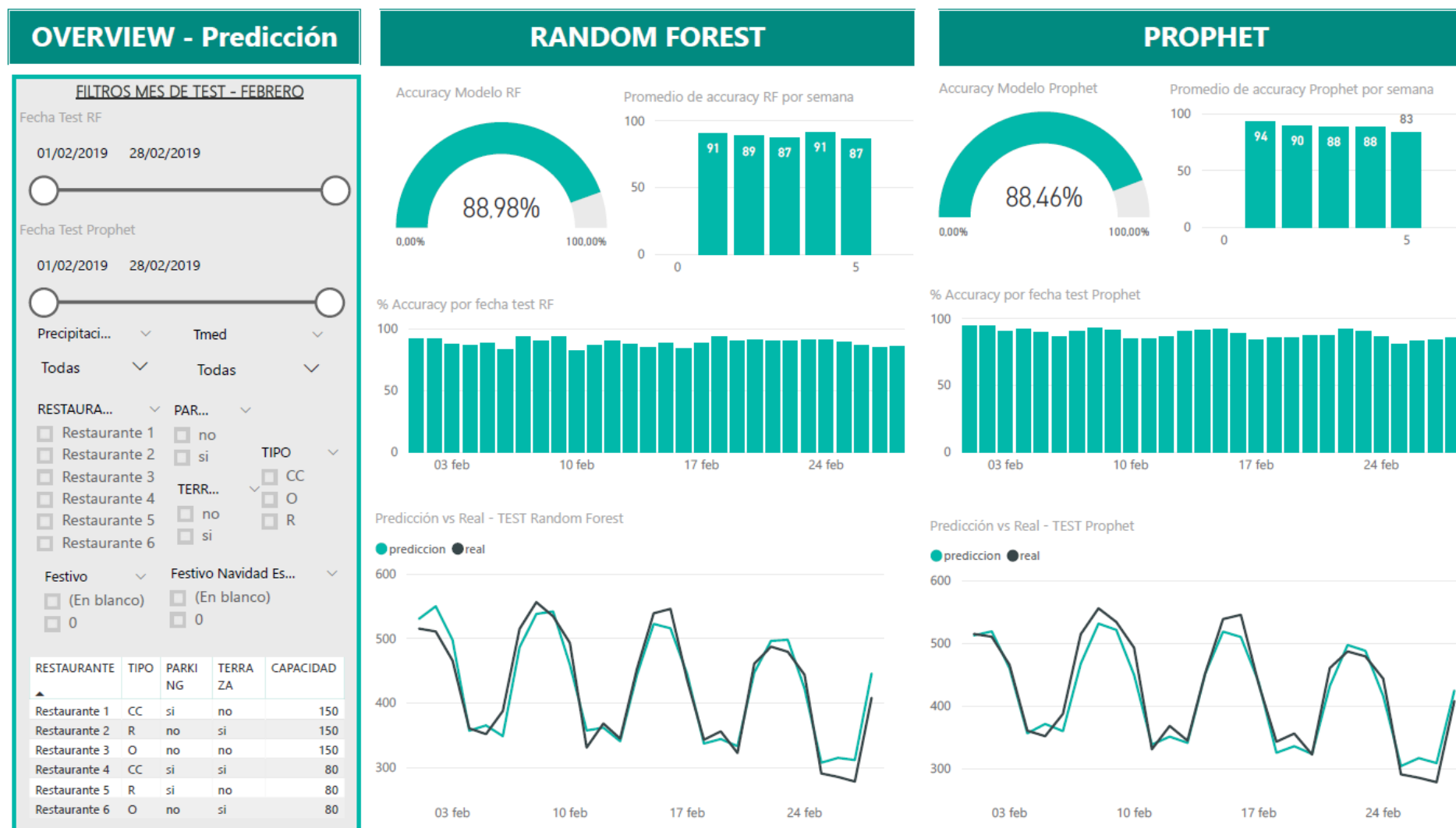


Figura 36. Segunda interfaz de la herramienta (Predicción)



Figura 37. Tercera interfaz de la herramienta (Optimización empleados)

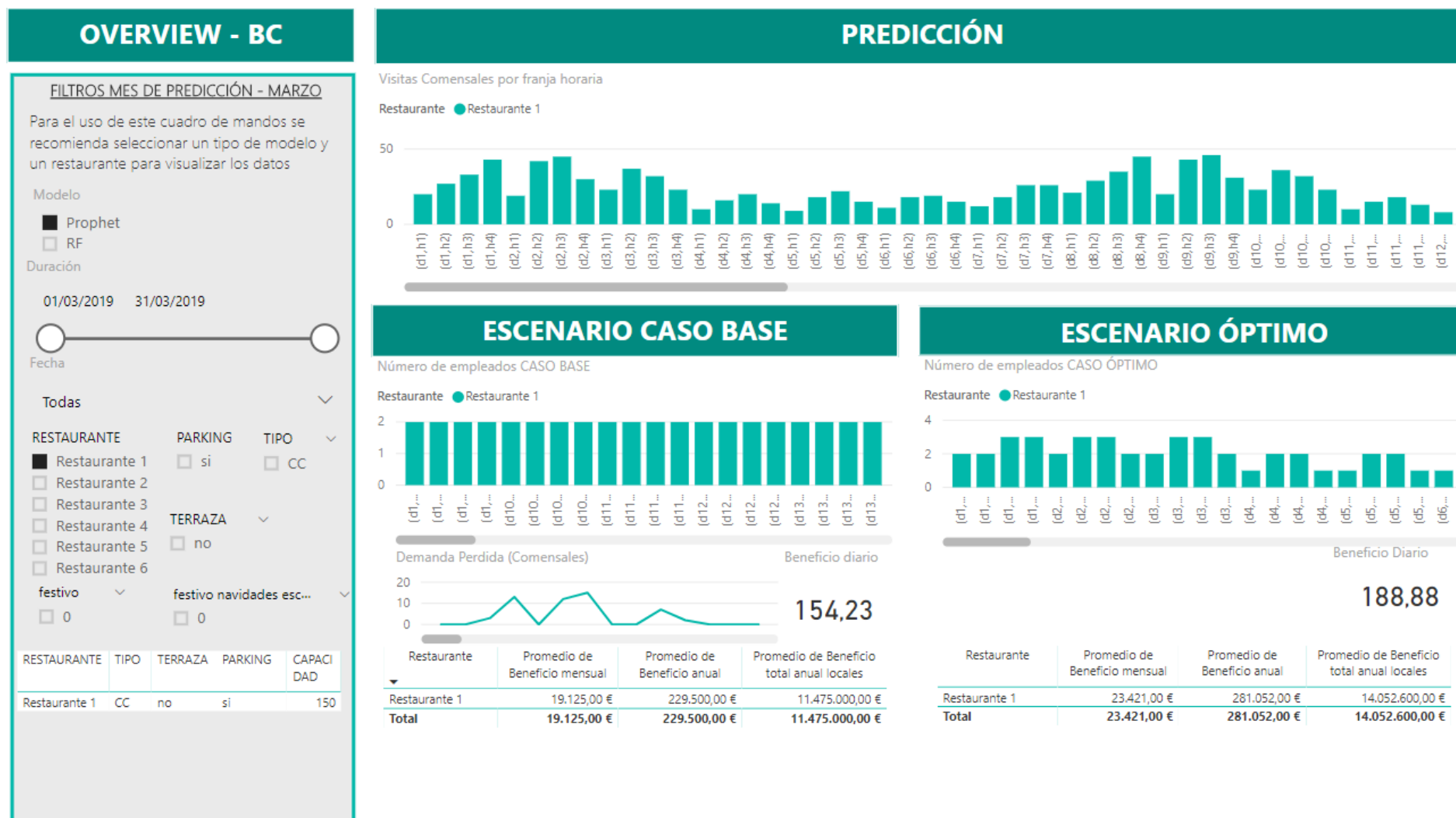


Figura 38. Tercera interfaz de la herramienta aplicando filtros

CAPÍTULO V : CASO DE NEGOCIO



5. Caso de negocio

5.1 Introducción

Después de haber conseguido optimizar la gestión del personal a nivel teórico es preciso analizar cuál sería el efecto e impacto de poner en práctica los resultados, de forma que no se quede en algo meramente hipotético e ideal. Para analizar el impacto de la herramienta se van a estudiar una serie de escenarios en los que se comparará el funcionamiento que tendrían los restaurantes simulados que componen la cadena de restauración tras aplicar la optimización respecto al modelo de funcionamiento tradicional utilizado en el presente. A continuación se detallará cuáles son los beneficios del modelo desde dos puntos de vista: a nivel económico y a nivel organizacional. Ello mostrará la eficacia y viabilidad del proyecto y el interés que puede aportar a potenciales clientes.

5.2 Escenarios e hipótesis

En primer lugar es necesario definir los escenarios que se van a estudiar y comparar para analizar el impacto:

- *Escenario 1:* El caso base corresponde a la situación en la cuál se contratarán dos empleados para locales con capacidad de 150 comensales (Restaurante 1, 2 y 3) y un empleado para los locales con capacidad de 80 comensales (Restaurante 4, 5 y 6). Del mismo modo, se mantiene la hipótesis de que el ratio de comensales que cada empleado atiende por hora es de 15. Este caso correspondería con el modo de funcionamiento actual de la cadena de restauración, sin aplicar ninguna optimización.
- *Escenario 2:* El siguiente escenario utilizará los empleados que la herramienta determina que son óptimos en cada franja horaria. Esto significa que el número de empleados podrá variar de una hora a la siguiente, durante las 4 horas de turno que ofrece el restaurante. El ratio de atención de comensales por parte de los empleados seguirá siendo de 15 personas.
- *Escenario 3:* Otro escenario planteado consiste en emplear durante todo el turno de 4 horas el mismo número de empleados, siendo este número el valor máximo obtenido en la optimización para dicho día.
- *Escenario 4:* Por último, se estudiará el efecto de ampliar el ratio de comensales atendidos por cada empleado a 20 personas en lugar de 15. Esto aumentará la carga de trabajo de los empleados.
-

Respecto a las hipótesis empleadas para calcular el beneficio se han mantenido aquellas planteadas en la optimización, que se resumen en los siguientes valores:

Precio venta	25,00 €/comensal
Coste empleado	11,00 €/hora
Coste pérdida	18,75 €/comensal perdido
Margen venta	9,38 €/comensal

Tabla 10. Hipótesis de costes asumidos para calcular beneficios

Puesto que en el modelo tan solo se han simulado seis restaurantes por simplificar los tiempos de ejecución del programa, es ahora el momento de escalar los resultados para que puedan simular una cadena de restauración más real. Es por ello que los resultados económicos obtenidos para seis restaurantes se escalarán multiplicando los resultados de cada tipo de restaurante por

10. En total, el número de restaurantes que formará la cadena será de 60, siendo 10 restaurantes similares al Restaurante 1, otros 10 restaurantes similares al Restaurante 2 y así sucesivamente. Además, en este caso el modelo de predicción solo ha calculado los resultados de predicción para un mes, de modo que para calcular el beneficio estimado anual se multiplicarán los resultados del mes calculado por 12 con el objetivo de obtener un resultado anual orientativo.

5.3 Estudio económico

En primer lugar se muestra el caso base que servirá de referencia para comparar los modelos optimizados y analizar las mejoras que generarían. Las tablas donde se muestran los resultados tienen tres columnas, la primera muestra el beneficio mensual obtenido para dicho restaurante, la segunda muestra el beneficio anual obtenido para ese restaurante y la última corresponde con el resultado que tendrían 10 locales del mismo tipo en un año. Por otro lado, se han analizado los resultados para ambos modelos de predicción, tanto el Random Forest como el Prophet.

La fórmula empleada para calcular los beneficios es:

$$\text{Beneficios} = N^{\circ}\text{comensales} * \text{Margen venta} - \text{Coste empleados} - \text{Coste pérdida de comensales}$$

Escenario 1:

Corresponde con el caso base, modelo que representa el funcionamiento actual de los locales. En este caso se mantiene el ratio de 15 comensales por empleado. Los beneficios obtenidos serían:

Modelo	Restaurante	Beneficio mensual	Beneficio anual	Beneficio total anual locales
RF	Restaurante 1	18.713 €	224.556 €	2.245.560 €
RF	Restaurante 2	17.747 €	212.964 €	2.129.640 €
RF	Restaurante 3	18.375 €	220.500 €	2.205.000 €
RF	Restaurante 4	9.417 €	113.004 €	1.130.040 €
RF	Restaurante 5	8.086 €	97.032 €	970.320 €
RF	Restaurante 6	9.145 €	109.740 €	1.097.400 €
Prophet	Restaurante 1	19.125 €	229.500 €	2.295.000 €
Prophet	Restaurante 2	17.325 €	207.900 €	2.079.000 €
Prophet	Restaurante 3	18.356 €	220.272 €	2.202.720 €
Prophet	Restaurante 4	9.914 €	118.968 €	1.189.680 €
Prophet	Restaurante 5	8.630 €	103.560 €	1.035.600 €
Prophet	Restaurante 6	9.089 €	109.068 €	1.090.680 €

Tabla 11. Beneficios caso base con ratio de 15 comensales por empleado

En total, para los resultados procedentes del Random Forest, la cadena de restauración obtendrá unos beneficios de 9.777.960,00 €. Para el modelo Prophet los resultados totales serán de 9.892.680,00 €.

Se muestra también a continuación cuál sería el caso base si se empleara un ratio de 20 comensales por empleado en lugar de 15, para poder compararlo con el escenario 4 posteriormente. El beneficio que se obtendrá será:

Modelo	Restaurante	Beneficio mensual	Beneficio anual	Beneficio total anual locales
RF	Restaurante 1	23.288 €	279.456 €	2.794.560 €
RF	Restaurante 2	22.716 €	272.592 €	2.725.920 €
RF	Restaurante 3	22.706 €	272.472 €	2.724.720 €
RF	Restaurante 4	12.267 €	147.204 €	1.472.040 €
RF	Restaurante 5	12.099 €	145.188 €	1.451.880 €
RF	Restaurante 6	11.920 €	143.040 €	1.430.400 €
Prophet	Restaurante 1	22.706 €	272.472 €	2.724.720 €
Prophet	Restaurante 2	24.488 €	293.856 €	2.938.560 €
Prophet	Restaurante 3	22.969 €	275.628 €	2.756.280 €
Prophet	Restaurante 4	12.352 €	148.224 €	1.482.240 €
Prophet	Restaurante 5	12.005 €	144.060 €	1.440.600 €
Prophet	Restaurante 6	11.545 €	138.540 €	1.385.400 €

Tabla 12. Beneficios caso base con ratio de 20 comensales por empleado

En este caso, para los resultados procedentes del Random Forest, la cadena de restauración obtendrá unos beneficios totales de 12.599.520,00 €. Para el modelo Prophet los resultados totales serán de 12.727.800,00 €. Se observa de esta forma que al aumentar el ratio de comensales que atenderá cada camarero se cubrirá la demanda antes y con menos camareros que en el caso previo, de modo que el coste de empleados disminuirá y el beneficio aumentará. En concreto, el beneficio ha aumentado en alrededor de un 28,86 % para el caso de Random Forest y un 28,65 % para el caso de Prophet. No obstante, en este caso la carga de trabajo de los empleados será mayor y podrá disminuir la calidad de las condiciones de trabajo del mismo.

Escenario 2:

En este caso se ha asumido que el número de empleados que van a trabajar en cada hora es el que la herramienta ha considerado como óptimo. De esta forma es posible que entre una hora y la siguiente haya una variación en el número de empleados. Si este modelo se llevara a cabo el beneficio obtenido para la cadena de restauración sería el siguiente:

Modelo	Restaurante	Beneficio mensual	Beneficio anual	Beneficio total anual locales
RF	Restaurante 1	25.409 €	304.908 €	3.049.080 €
RF	Restaurante 2	24.001 €	288.012 €	2.880.120 €
RF	Restaurante 3	24.279 €	291.348 €	2.913.480 €
RF	Restaurante 4	14.044 €	168.528 €	1.685.280 €
RF	Restaurante 5	14.178 €	170.136 €	1.701.360 €
RF	Restaurante 6	12.718 €	152.616 €	1.526.160 €
Prophet	Restaurante 1	23.421 €	281.052 €	2.810.520 €
Prophet	Restaurante 2	27.595 €	331.140 €	3.311.400 €
Prophet	Restaurante 3	24.636 €	295.632 €	2.956.320 €
Prophet	Restaurante 4	13.096 €	157.152 €	1.571.520 €
Prophet	Restaurante 5	13.126 €	157.512 €	1.575.120 €
Prophet	Restaurante 6	12.252 €	147.024 €	1.470.240 €

Tabla 13. Beneficios caso óptimo con ratio 15 comensales por empleado

Se observan en los resultados que el beneficio obtenido es superior al del caso base. En concreto, anualmente la cadena de restauración obtendrá un total de 13.755.480,00 € según el modelo Random Forest, y un total de 13.695.120,00 € según el modelo Prophet. Esto supondrá un incremento del 40,67 % de los beneficios de la cadena utilizando el modelo Random Forest, con un ratio de 15 comensales por empleado, y un incremento del 38,44% en el modelo Prophet.

Escenario 3:

Este escenario refleja cuál sería la situación económica de la cadena si decidieran emplear en el turno de servicio completo el mismo número de empleados, siendo este valor el número máximo de empleados que la herramienta ha calculado que será necesaria en ese día durante algún periodo.

Modelo	Restaurante	Beneficio mensual	Beneficio anual	Beneficio total anual locales
RF	Restaurante 1	24.638,88 €	295.666,50 €	2.956.665,00 €
RF	Restaurante 2	23.230,50 €	278.766,00 €	2.787.660,00 €
RF	Restaurante 3	23.080,38 €	276.964,50 €	2.769.645,00 €
RF	Restaurante 4	13.769,25 €	165.231,00 €	1.652.310,00 €
RF	Restaurante 5	13.881,00 €	166.572,00 €	1.665.720,00 €
RF	Restaurante 6	12.233,88 €	146.806,50 €	1.468.065,00 €
Prophet	Restaurante 1	22.639,88 €	271.678,50 €	2.716.785,00 €
Prophet	Restaurante 2	26.605,38 €	319.264,50 €	3.192.645,00 €
Prophet	Restaurante 3	23.436,63 €	281.239,50 €	2.812.395,00 €
Prophet	Restaurante 4	12.810,13 €	153.721,50 €	1.537.215,00 €
Prophet	Restaurante 5	12.762,50 €	153.150,00 €	1.531.500,00 €
Prophet	Restaurante 6	11.757,38 €	141.088,50 €	1.410.885,00 €

Tabla 14. Beneficios caso óptimo redondeando el número de empleados

Según los resultados obtenidos para este escenario, para el modelo Random Forest se obtiene un total de beneficios para la cadena de 13.300.065,00 €, y para el modelo Prophet se obtienen 13.201.425,00 €. Esto supone un valor ligeramente inferior al obtenido en el Escenario 2, en concreto supone una diferencia de 455.415,00 € menos en este caso para el modelo Random Forest y 493.695,00 € para el modelo Prophet. No obstante, para ambos modelos el resultado obtenido es muy superior al beneficio correspondiente al Escenario 1. De modo que la optimización realizada en este escenario mejoraría también los resultados del caso base. En este caso el incremento sería de 36 % para el modelo Random Forest y 33,45 % para el modelo Prophet.

Escenario 4:

En este último escenario lo que se ha realizado ha sido un cambio en el ratio de comensales que atenderá cada empleado. En este caso se va a fijar ese valor a 20 comensales por trabajador. Se va a asumir que se emplearán los trabajadores concretos que especifica la herramienta en cada hora, de forma que habrá variación horaria entre el número de empleados atendiendo. En este caso el beneficio de cada restaurante será el siguiente:

Modelo	Restaurante	Beneficio mensual	Beneficio anual	Beneficio total anual locales
RF	Restaurante 1	26.014 €	312.168 €	3.121.680 €
RF	Restaurante 2	24.507 €	294.084 €	2.940.840 €
RF	Restaurante 3	24.862 €	298.344 €	2.983.440 €
RF	Restaurante 4	14.275 €	171.300 €	1.713.000 €
RF	Restaurante 5	14.398 €	172.776 €	1.727.760 €
RF	Restaurante 6	13.026 €	156.312 €	1.563.120 €
Prophet	Restaurante 1	23.916 €	286.992 €	2.869.920 €
Prophet	Restaurante 2	28.211 €	338.532 €	3.385.320 €
Prophet	Restaurante 3	25.230 €	302.760 €	3.027.600 €
Prophet	Restaurante 4	13.349 €	160.188 €	1.601.880 €
Prophet	Restaurante 5	13.357 €	160.284 €	1.602.840 €
Prophet	Restaurante 6	12.516 €	150.192 €	1.501.920 €

Tabla 15. Beneficios caso óptimo aumentando el ratio de comensales por empleado a 20.

En este caso se obtendrá un total de 14.049.840,00 € de beneficios para la cadena de restauración según el modelo Random Forest, y un total de 13.989.480,00 € según el modelo Prophet. Este resultado es incluso mayor que los obtenidos en el resto de escenarios ya que coincide con la optimización máxima posible al estar realizada por horas adaptando el número de empleados en cada hora del turno, y además ya que se ha aumentado el ratio de clientes que va a poder atender cada empleado. No obstante la carga de trabajo de los mismos aumentará también y la calidad de servicio se podrá ver afectada negativamente. En concreto, se incrementan los beneficios respecto al caso base en un 43.69 % para el modelo Random Forest y en un 41.41 % para el modelo Prophet.

5.4 Impacto y análisis de los beneficios

Una vez conocidos los resultados de los distintos escenarios se procede a analizar cuál sería el impacto de la implantación de los mismos. Este impacto se puede analizar desde dos puntos de vista, el económico y el organizacional.

5.4.1 Nivel económico

En las siguientes tablas se muestra un resumen de los resultados obtenidos por restaurantes para cada uno de los escenarios. Esto nos permite observar rápidamente cuál será el mejor de los casos. Observando únicamente los números es posible identificar que el escenario más beneficioso en cuanto a resultados económicos es el escenario 4. Este escenario resultaba del incremento en el ratio del número de comensales de 15 a 20. Esto significa que para atender la misma demanda necesitaremos menos empleados que anteriormente, lo cual disminuirá los costes de los mismos. No obstante, como ya se ha mencionado previamente, la calidad del servicio ofrecido podrá disminuir ya que los empleados tendrán más carga de trabajo, pudiendo dedicar a cada comensal menos tiempo del que se dedicaba previamente. Además, al igual que el escenario 2, el escenario 4 ajusta horariamente el número de empleados, lo cual hace que incrementen los resultados pero que se complique la organización y gestión del personal. Se podría aplicar lo mismo que en el escenario 3 donde se redondea el número de empleados para utilizar el mismo número durante todo el turno, siendo este el máximo necesitado en algún momento del turno. Al realizar este redondeo los resultados económicos disminuirían ligeramente ya que se contratarán más empleados que en el caso óptimo. Este efecto es similar al ocurrido entre el escenario 2 y 3.

Por otro lado, la diferencia económica entre el escenario 4 y el escenario 2 no es demasiado elevada. Por lo que el siguiente caso más beneficioso para la empresa numéricamente hablando sería el escenario 2. En este caso se mantiene el ratio de calidad establecida para la cadena de restauración en 15 comensales por empleado. Lo que ocurre en este escenario es que se adapta el número de empleados necesarios por cada hora, es decir, el número podrá variar entre una hora y la siguiente. Es por ello que los resultados numéricos obtenidos son elevados, ya que se está adaptando el negocio a las necesidades de los clientes en tiempo real. Sin embargo, este escenario puede ser complicado de llevar a cabo como se explicará en el siguiente apartado.

Si se observan los resultados del escenario 3, la diferencia con el escenario 2 tampoco es excesiva. Ciertamente es que los ingresos son menores, pero a nivel de organización puede resultar un escenario interesante a considerar. En este caso, se utilizan los resultados de número de empleados necesarios en el caso óptimo y con un ratio de calidad de atención al cliente de 15 comensales, pero el número de empleados a lo largo de todo el turno no varía. Este número corresponde con el valor máximo de empleados obtenido para alguna de las horas del turno en cuestión. De esta forma, las necesidades de las horas punta estarán cubiertas sin problema, y durante las horas de menor afluencia los empleados que estarán menos cargados de trabajo podrán dedicar tiempo a otras tareas como por ejemplo: personalización de la oferta para los clientes, búsqueda y atracción activa de nuevos comensales, tareas de organización del inventario, tareas de limpieza, mejora de la calidad de servicio y del local...

En cualquier caso, se observa que todos los escenarios de mejora propuestos incrementan sustancialmente los beneficios obtenidos para cada restaurante. Aún eligiendo el escenario que “menos” mejora los resultados del caso base (escenario 3), los resultados aumentan en un 36% para el Random Forest y un 33,45 % para el Prophet. Lo que significa, que en el peor de los casos esta herramienta supondría una gran ventaja competitiva a nuestros clientes frente a sus competidores al permitirles incrementar sus beneficios como mínimo en alrededor del 30%, simplemente optimizando la gestión de los empleados, lo cual tendrá bastante impacto a nivel económico.

El resumen de resultados mensuales para el Random Forest se muestra a continuación:

BENEFICIOS	ESCENARIO 1: CASO BASE RATIO 15	ESCENARIO 2: CASO ÓPTIMO EXACTO	ESCENARIO 3: CASO REDONDEO	ESCENARIO 4: RATIO 20 - EXACTO
Restaurante 1	18.713 €	25.409 €	24.639 €	26.014 €
Restaurante 2	17.747 €	24.047 €	23.231 €	24.507 €
Restaurante 3	18.375 €	24.279 €	23.080 €	24.862 €
Restaurante 4	9.417 €	14.044 €	13.769 €	14.275 €
Restaurante 5	8.086 €	14.178 €	13.881 €	14.398 €
Restaurante 6	9.145 €	12.718 €	12.234 €	13.026 €

Tabla 16. Resumen beneficios mensuales de cada escenario para datos del Random Forest

Los resultados en el caso del modelo Prophet serían:

BENEFICIOS	ESCENARIO 1: CASO BASE RATIO 15	ESCENARIO 2: CASO ÓPTIMO EXACTO	ESCENARIO 3: CASO REDONDEO	ESCENARIO 4: RATIO 20 - EXACTO
Restaurante 1	19.125 €	23.421 €	22.640 €	23.916 €
Restaurante 2	17.325 €	27.595 €	26.605 €	28.211 €
Restaurante 3	18.356 €	24.636 €	23.437 €	25.230 €
Restaurante 4	9.914 €	13.096 €	12.810 €	13.349 €
Restaurante 5	8.630 €	13.126 €	12.763 €	13.357 €
Restaurante 6	9.089 €	12.252 €	11.757 €	12.516 €

Tabla 17. Resumen beneficios mensuales de cada escenario para datos del Random Forest

5.4.2 Nivel organizacional

A nivel de organización, esta herramienta puede suponer también un gran cambio y tener consecuencias muy positivas a este respecto. Las ideas propuestas en este apartado deberían ser revisadas por un equipo de legislación laboral para verificar su validez y capacidad de aplicación.

En primer lugar, puesto que esta herramienta está destinada a cadenas de restauración que contarán con distintos locales pertenecientes a cada marca, se propone crear una bolsa de trabajo por cada marca dentro de la cadena de la que formen parte los empleados. Lo ideal sería que los empleados pudieran seleccionar una de las zonas o áreas de Madrid donde se sitúen una serie de restaurantes de la marca, de forma que el empleador tenga la posibilidad de destinarle a uno u otro centro en función de las necesidades. Puesto que la herramienta realiza una tarea de predicción, sería posible conocer las necesidades del restaurante con antelación, de forma que el empleado pueda ser avisado con tiempo y comprobar su disponibilidad. A pesar de que existe una movilidad para el empleado, esta estaría restringida a una zona determinada con un límite de extensión.

A nivel organizacional se estima que el caso más fácil de llevar a cabo obteniendo mejoras respecto al caso base sería el escenario 3. En este escenario no se producen variaciones entre el número de empleados de una hora a la siguiente, por lo que solo podría cambiar la localización del centro de trabajo de un día a otro. No obstante, si pudiera ponerse en práctica el escenario 2 o 4, se podrían movilizar los empleados entre unos locales y otros de la misma marca de unas horas a otras en función de las necesidades para cubrir la demanda dentro del mismo turno. En este caso se alcanzarían los mejores resultados económicos. Sin embargo, se consideran escenarios menos factibles ya que habría que contar con el tiempo de desplazamiento y el coste del mismo, y legalmente parece ser más complicado. Es por eso que respecto a la organización se preferiría el escenario 3.

Se ha demostrado en el apartado anterior el impacto que tendría en los resultados de la cadena el hecho de optimizar la gestión del personal. El personal es el activo más importante de las empresas y actualmente a penas se dedica tiempo a mejorar su gestión ya que no se había estimado el efecto que esto supondría.

Con esta herramienta no solo mejorarán los resultados de la empresa, también mejorará la calidad del servicio ofrecido al poder personalizar la carga de trabajo de los empleados, lo cuál incrementará el nivel de satisfacción de los clientes y en consecuencia mejorará la imagen de marca incrementando la fidelidad de los clientes. Se podrá también tener información real sobre el nivel de ocupación de cada empleado pudiendo optimizar su rendimiento y gestionar las tareas de manera más eficiente. Además, mejorará al mismo tiempo la calidad de trabajo de los empleados y se aumentará la lealtad a la empresa y sentimiento de pertenencia, reteniendo mejor el talento.

CAPÍTULO VI : RESULTADOS Y CONCLUSIONES



6. Resultados y conclusiones

6.1 Conclusiones del proyecto

Desde mi punto de vista, este proyecto supone una gran innovación y revolución en el sector de restauración que permite resolver un problema real detectado en este tipo de clientes. Además, se trata de una herramienta muy completa que incluye conocimientos multidisciplinarios y cuya aplicación podrá extenderse a otros sectores. Por otro lado, esta herramienta favorecerá al proceso de digitalización en el que está introduciéndose este sector, y facilitará el uso y la aplicación de algoritmos de aprendizaje supervisado e inteligencia artificial al proceso de gestión de empleados.

Los objetivos planteados al inicio del proyecto se han cumplido satisfactoriamente y se ha podido demostrar numéricamente la eficacia de la herramienta y el impacto que tendría su utilización en un cliente prototipo de una cadena de restauración. Además, se trata de una herramienta completa, flexible y fácil de adaptar a los clientes, pudiendo incorporar modificaciones y aspectos personalizados en cada caso.

Como se ha detallado en la memoria, esta herramienta consta principalmente de tres partes, que se pueden visualizar en cada una de las interfaces de la herramienta: el histórico e indicadores clave de rendimiento, la predicción de afluencia de comensales en restaurantes y la optimización del número de empleados necesarios. Cada interfaz es dinámica pudiendo seleccionar la información general que resume la cadena de restauración, o seleccionando un restaurante o local en particular dentro de la misma y observando tan solo esa información. Además, incluye una serie de filtros que permitirán personalizar también la información que se desea visualizar. Ello supone una ventaja de utilización, siendo una herramienta muy fácil de usar y que permite analizar resultados de manera simple y concreta. Todo ello permitirá extraer conclusiones de manera rápida para poder actuar en consecuencia según las necesidades y mejorar así el negocio en sus distintos aspectos.

La realización de este proyecto ha permitido profundizar en ciertos aspectos de conocimiento aprendidos durante el máster y las prácticas realizadas como por ejemplo modelos predictivos, procesos de optimización, operaciones, sistemas de organización... y se han podido conocer nuevos ámbitos de conocimiento, como por ejemplo los referentes al sector de restauración y las técnicas aplicadas en él. También ha permitido darse cuenta del impacto que las nuevas tecnologías pueden llegar a tener en los negocios, y las grandes mejoras que pueden aportar a ciertos sectores, colaborando en el desarrollo de los mismos. Además de facilitar la vida a los usuarios de estas herramientas, permite también centrar el negocio cada vez más en el cliente pudiendo anticiparse a sus necesidades y ofrecerle exactamente lo que necesita en cada momento. Esto supone en muchos casos una ventaja competitiva respecto a los competidores en el momento de lanzamiento de este tipo de aplicaciones o herramientas, y generalmente acaban suponiendo un elemento fundamental y prácticamente obligatorio para todos los participantes del sector una vez se extiende su uso si pretenden perdurar en el mismo.

Una de las dificultades encontradas para realizar este proyecto ha sido la ausencia de datos disponibles. Ello ha dificultado mucho el proceso de construcción de la herramienta y ha supuesto mucho esfuerzo y tiempo. No obstante, puesto que se confiaba en el gran impacto que esta herramienta podía tener en la realidad y que se contaba con un grupo de expertos en el sector importante, se decidió emplear datos simulados que pudieran reflejar la realidad de la manera más exacta posible. Esto ha podido influenciar en los datos obtenidos.

El siguiente paso y objetivo final de esta herramienta será el de incorporar una base de datos real de un cliente a su fuente de alimentación de forma que se pueda realizar la predicción y

optimización basándose en dichos datos. La herramienta está preparada para poder incorporar una base de datos real fácilmente, de forma que la instalación de la herramienta en el cliente será rápida y sencilla, y podrá usarse acto seguido.

Respecto a los resultados mostrados en la memoria presentada no siempre se han incluido todos los resultados obtenidos para cada restaurante, ya que se consideraba irrelevante en algunos casos y por evitar repeticiones en otros. Se ha pretendido mostrar los resultados que podían servir de referencia y/o ejemplo en cada caso, y las repeticiones realizadas para el resto de restaurantes no mostradas se han incluido al final de esta memoria, en el apartado de *Anexos*, por si se desea acceder a dicha información.

En cuanto a la metodología seguida para la realización del proyecto, se ha contado con un grupo de expertos de la firma PwC que han ido supervisando y controlando los resultados que se obtenían para cada uno de los componentes de la herramienta. En muchos casos los resultados mostrados son los obtenidos al final de un largo proceso en el cuál se han realizado múltiples modificaciones, pero se ha decidido evitar mencionarlas buscando siempre la mayor precisión y concisión posible.

Finalmente, se han detectado una serie de mejoras del proyecto que podrían realizarse de cara a vender esta herramienta a un cliente real o de cara a introducirla en el mercado, que se mostrarán a continuación.

Para concluir, se considera que la herramienta tiene mucho potencial dentro del sector y podrá generar un gran impacto suponiendo muchas mejoras para los restaurantes, solucionando los problemas existentes a día de hoy e implicando un paso adelante en la gestión del personal, conllevando mejoras a nivel global dentro del negocio, tanto de organización y previsión como económicas.

6.2 Próximos pasos y mejoras

A continuación se proponen distintas medidas con las que mejorar la herramienta creada buscando el perfeccionamiento de la misma de cara a una posible venta a un cliente real.

6.2.1 Simulación

El apartado de simulación del funcionamiento de una cadena de restauración es simplemente teórico y no sería necesario si se contara con una base de datos real de una cadena de restauración. En ese caso, la base de datos serviría para extraer los patrones de comportamiento directamente, y se utilizaría para entrenar los modelos de predicción.

No obstante, respecto a la simulación realizada, podría mejorarse si se continuara con el procedimiento seguido del Método Delphi. Este proceso iterativo decidió finalizarse por cuestión de tiempo una vez se tenían resultados aceptables según los expertos que reflejaban la realidad, pero los parámetros podrían seguir refinándose si se realizaran más iteraciones.

Por otro lado también se podría mejorar la simulación incorporando nuevas reglas de generación que afectaran a los factores y añadieran más precisión a la misma.

6.2.2 Predicción

Respecto a la predicción, además de los modelos matemáticos empleados se podría incorporar algún modelo más como por ejemplo el ARIMA, o redes neuronales, con la intención de comprobar cuál ofrece mejores resultados.

Sería posible también continuar refinando los parámetros empleados en los modelos predictivos o añadir nuevas variables que pudieran incrementar la precisión de la predicción.

En concreto, para el modelo Random Forest, una vez averiguadas las importancias de las matrices sería posible volver a ejecutar la predicción eliminando aquellas matrices cuya importancia sea nula. De esta forma se espera obtener resultados más precisos.

Para el modelo Prophet, podrían incluirse más eventos anuales que puedan afectar al modelo y la afluencia como por ejemplo: fechas de partidos de fútbol, baloncesto u otros deportes, conciertos, eventos de empresas... También podrían incorporarse en el modelo ofertas que se realicen en los restaurantes, para incluir dicha variable en la predicción y poder analizar cómo afectaría.

6.2.3 Optimización

La optimización de empleados podría mejorarse personalizando a los empleados en la herramienta. Es decir, se podrían clasificar según su categoría (camareros, cocineros, jefes de servicio, personal de limpieza...) y que el modelo sea capaz de calcular cuántos empleados de cada tipo son necesarios.

Además, sería buena idea tener la posibilidad de añadir información personal de cada empleado incluyendo las disponibilidades o restricciones de cada uno de forma que además de indicar el número de empleados, indique qué persona tiene que realizar cada turno.

6.3 Posibles aplicaciones

Esta herramienta puede tener distintas aplicaciones además de la mencionada a lo largo del proyecto. En primer lugar, puede aplicarse al sector de consultoría, permitiendo hacer proyectos con cadenas de restauración y ofrecerles a las mismas un valor diferencial. Además, en muchos casos de consultoría ocurre que al iniciar un proyecto o una propuesta de proyecto no se dispone de la información necesaria sobre el negocio como para realizar proposiciones. Esta herramienta, y en concreto la parte de simulación, podría ayudar a estos proyectos aportando una simulación del negocio para poder analizar cómo se va a comportar y qué opciones tiene la consultora para recomendar a su cliente.

Otra posible aplicación podría ser para clientes que deseen abrir un nuevo negocio de restauración pero que no dispongan de mucha información ni experiencia en el sector. Esta herramienta les permitiría analizar una simulación del mismo y ver cómo afectarían distintas variables en su negocio. Al mismo tiempo les permitiría analizar qué localización es óptima para su negocio.

Por último, puesto que se trata de una herramienta que optimiza la gestión del personal sería posible aplicar esta herramienta a diversos sectores además del de restauración, como por ejemplo: tiendas y comercios, hoteles, museos, discotecas, caterings, parques recreativos, servicios de bomberos e incluso clínicas hospitalarias o centros médicos. No obstante, para aplicarlos a estos sectores es probable que haya que modificar los modelos predictivos y adaptar las variables a aquellas que más puedan afectar al sector o negocio en cuestión. En principio esto no sería muy complicado, bastaría con analizar el negocio en profundidad y extraer patrones de comportamiento y factores de influencia en el mismo.

CAPÍTULO VII : BIBLIOGRAFÍA



7. Bibliografía

- [BOIZ18] Boiza, G. (2018). Los bares y restaurantes españoles crecen por tercer año, con un aumento de ventas del 2,5% en 2017. Retrieved from <https://www.eleconomista.es/distribucion/noticias/8947691/02/18/Los-bares-y-restaurantes-espanoles-crecen-por-tercer-ano-con-un-aumento-de-ventas-del-25-en-2017-.html>
- [BREI18] Breiman, L., & Cutler, A. (2018). Random Forest.
- [BROWND] Brown, E. D. (n.d.). Forecasting Time-Series data with Prophet – Part 1. Retrieved from <https://pythondata.com/forecasting-time-series-data-with-prophet-part-1/>
- [CAIXND] Caixa Bank. (n.d.). *Manual de gestión de un restaurante*. Retrieved from <https://www.caixabanklab.com/elbullifoundation/manual-gestion-restaurant/>
- [CANA18] CANADEAN. (2020). *Spain Foodservice : The Future of Foodservice in Spain to 2020*.
- [CASA18] Casares, J., & Catedr, R. (2018). *Decálogo quintaesenciado de cuestiones relevantes sobre comer y beber fuera del hogar* (Vol. 4).
- [CHOU18] CHOUDHARY, A. (2018). Generate Quick and Accurate Time Series Forecasts using Facebook’s Prophet (with Python & R codes). Retrieved from <https://www.analyticsvidhya.com/blog/2018/05/generate-accurate-forecasts-facebook-prophet-python-r/>
- [CICEND] Cicero Comunicacion. (n.d.). Método Delphi. Retrieved from <https://www.cicerocomunicacion.es/que-es-el-metodo-delphi/>
- [COSSI18] Cossío, A. (2018). Bots, Machine Learning, Servicios Cognitivos Realidad y perspectivas de la Inteligencia Artificial en España, 2018. In *Pwc*. Retrieved from <https://www.pwc.es/es/publicaciones/tecnologia/assets/pwc-ia-en-espana-2018.pdf>
- [CUIN18] Cuiner. (2018). Cuiner Software de restauración. Retrieved from <https://cuiner.com/software-modulos-opcionales/quierofactura/>
- [DBKI17] DBK Informa. (2017). *Informes Especiales Cadenas de restaurantes*.
- [ORGAND] Departamento de Organización Industrial. (n.d.). *Quality Management 1819 Moodle version*.
- [ECHE17] Echegaray, J. G. (2017). El crecimiento del PIB se produce con una mínima mejora de la productividad. Retrieved from <https://www.libremercado.com/2017-09-03/el-crecimiento-del-pib-se-produce-con-una-minima-mejora-de-la-productividad-1276605177/>
- [EURO17] Europa Press. (2017). La restauración española consolida su recuperación tras crecer un 2% en 2016. Retrieved from

<https://www.eleconomista.es/economia/noticias/8140721/02/17/Economia-La-restauracion-espanola-consolida-su-recuperacion-tras-crecer-un-2-en-2016.html>

- [EXPA17] Expansión. (2017). *Más digitales e inteligentes: Así serán los restaurantes del futuro*. Retrieved from <http://www.expansion.com/economia-digital/innovacion/2017/02/15/589ca279e5fdea0a358b4596.html>
- [FACEND] Facebook. (n.d.). Non daily data - Prophet. Retrieved from https://facebook.github.io/prophet/docs/non-daily_data.html
- [FACEND] Facebook. (n.d.). Quick Start Prophet. Retrieved from https://facebook.github.io/prophet/docs/quick_start.html#python-api
- [FIGU17] Figuerola Palomo, M., Sánchez Polo, M., & Herranz Guillén, J. L. (2017). *Estudio Anual del Sector de la Hostelería en España*. Retrieved from www.fehr.es
- [GALL18] Gallego, E. (2018). *El sector de la restauración en España* (Vol. 4). Retrieved from www.fehr.es
- [GARC17] García, P. M. (2017). *¿Cuánto personal necesita tu restaurante?* Retrieved from <https://enlacocina.telemesa.es/gestion-administracion-restaurantes/cuanto-personal-necesitas-restaurant/>
- [GARC17] García, P. M. (2017). *LOS MEJORES SOFTWARE DE GESTIÓN PARA RESTAURANTES*. Retrieved from <https://enlacocina.telemesa.es/actualidad/los-mejores-software-gestion-restaurantes/>
- [CARG18] García Ponce, M. (2018). *MODELOS ANALÍTICOS E INTELIGENCIA ARTIFICIAL PARA LA MEJORA DE OPERACIONES*. Universidad Pontificia de Comillas, ICAI.
- [CARG18] García, P. M. (2018). *LA INTELIGENCIA ARTIFICIAL REVOLUCIONA EL MUNDO DE LA HOSTELERÍA*. Retrieved from <https://enlacocina.telemesa.es/gestion-administracion-restaurantes/la-inteligencia-artificial-revoluciona-el-mundo-de-la-hosteleria/>
- [GURO18] Gurobi. (n.d.). The Gurobi Python Interface for Python Users. Retrieved from https://www.gurobi.com/documentation/8.1/quickstart_windows/the_gurobi_python_interface.html
- [HORE17] Horeca Profesional. (2017). La restauración española crece un 4,7% más en 2017 y encadena cuatro años al alza. Retrieved from <http://profesionalthoreca.com/la-restauracion-espanola-crece-47-mas-2017-encadena-cuatro-anos-al-alza/>
- [HOST18] Hostelería Digital. (2018). Importante salto de la digitalización en los restaurantes y hoteles. Retrieved from <https://www.hosteleriadigital.es/2018/02/21/importante-salto-de-la-digitalizacion-en-los-restaurantes-y-hoteles/>

- [HOST15] Hosteltur. (2015). La facturación de los restaurantes detiene su caída tras siete años de crisis. Retrieved from https://www.hosteltur.com/195530_facturacion-restaurantes-detiene-su-caida-siete-anos-crisis.html
- [IFMA17] IFMA. *Grandes cifras del FM restauración 2017 4ª Edición.* , (2017).
- [KOEH17] Koehrsen, W. (2017). Random Forest in Python. Retrieved from <https://towardsdatascience.com/random-forest-in-python-24d0893d51c0>
- [KOEH18] Koehrsen, W. (2018). Improving Random Forest in Python Part I. Retrieved from <https://towardsdatascience.com/improving-random-forest-in-python-part-1-893916666cd>
- [KPMG18] KPMG. (2018). La restauración se alimenta con datos. Retrieved from <https://www.tendencias.kpmg.es/2018/09/restauracion-big-data/>
- [KPMG18] KPMG. (2018). 4 | *Anuario de la restauración en España*. Retrieved from <https://assets.kpmg/content/dam/kpmg/es/pdf/2018/09/anuario-restauracion-organizada-espana.pdf>
- [LOPE18] López, C. (2018, September 13). Las cadenas de restauración elevan su perspectiva de crecimiento para 2018 por el tirón de la economía. *El País*. Retrieved from https://elpais.com/economia/2018/09/12/actualidad/1536761748_703168.html
- [MAAM18] Ma'amari, M. (2018). Introduction to Random Forest Algorithm with Python. Retrieved from <https://blog.goodaudience.com/introduction-to-random-forest-algorithm-with-python-9efd1d8f0157>
- [MALI18] Malik, U. (2018). Random Forest Algorithm with Python and Scikit-Learn. Retrieved from <https://stackabuse.com/random-forest-algorithm-with-python-and-scikit-learn/>
- [MEEK18] Meek, A. (2018). *How Technology Is Disrupting the Multibillion Dollar Meetings Industry*. 1–9. Retrieved from https://cumoodle.coventry.ac.uk/pluginfile.php/2617762/mod_resource/content/0/EventIndustry_How Technology is Disrupting the Multibillion Dollar Meetings Industry.pdf
- [MILE19] Miles, S. (2019). *5 Predictive Ordering Platforms for Restaurants*.
- [NPDGND] NDP Group. (n.d.). Los bares y restaurantes españoles crecen por tercer año, con un aumento de las ventas del 2,5% en 2017. Retrieved from <https://www.npdgroup.es/wps/portal/npd/es/noticias/comunicados-de-prensa/los-bares-y-restaurantes-espanoles-crecen-por-tercer-ano-con-un-aumento-de-las-ventas-del-2-5-en-2017/>
- [NPDG18] Ndp Group. (2018). Los bares y restaurantes españoles crecen por tercer año, con un aumento de las ventas del 2,5% en 2017. Retrieved from <https://www.npdgroup.es/wps/portal/npd/es/noticias/comunicados-de-prensa/los-bares-y-restaurantes-espanoles-crecen-por-tercer-ano-con-un-aumento-de-las-ventas-del-2-5-en-2017/>

- [NIET17] Nieto, M. Á. (2017). Prophet: Forecasting our Metrics (or Predicting the Future). Retrieved from <https://www.percona.com/blog/2017/03/20/prophet-forecasting-our-metrics-or-predicting-the-future/>
- [OREL18] Orellana Alvear, J. (2018). Parámetros árboles de decisión. Retrieved from <https://bookdown.org/content/2031/arboles-de-decision-parte-ii.html>
- [ORQU18] Orquest. (2018). Por qué usar Excel en la planificación de tu restaurante te está haciendo perder rentabilidad. Retrieved from <https://orquest.es/usar-excel-en-planificacion-restaurante-hace-perder-rentabilidad/>
- [RUIZ18] Ruiz, A. (2018). Find out how artificial intelligence applied to restoration works. Retrieved from <https://www.diegocoquillat.com/en/descubre-como-funciona-la-inteligencia-artificial-aplicada-a-la-restauracion/>
- [SAMA16] Samani, M. (2016). *Industry Perspectives Restaurant Sector Changing Dynamics of Restaurant Industry*.
- [SEAN17] Sean J, T., & Letham, B. (2017). Prophet: forecasting at scale. Retrieved from <https://research.fb.com/prophet-forecasting-at-scale/>
- [SENN18] Sennaar, K. (2018). Examples of AI in Restaurants and Food Services Restaurant Chatbots / Conversational Interfaces Say2eat. Retrieved from <https://www.techemergence.com/ai-in-restaurants-food-services/>
- [SING18] SINGH, A. (2018). Machine Learning Part I. Retrieved from <https://www.analyticsvidhya.com/blog/2018/10/comprehensive-overview-machine-learning-part-1/>
- [SING18] SINGH, A. (2018). Random Forest Interpretation. Retrieved from <https://www.analyticsvidhya.com/blog/2018/10/interpret-random-forest-model-machine-learning-programmers/>
- [SOLDND] Soldevila Matías, P. (n.d.). *Recursos Humanos y Dirección de equipos en restauración*.
- [TAYL18] Taylor, S. J., Letham, B., Taylor, S. J., & Letham, B. (2018). Forecasting at Scale (Prophet). In *The American Statistician*. <https://doi.org/10.1080/00031305.2017.1380080>
- [TAYL17] Taylor, S. J., & Letham, B. (2017). *Forecasting at Scale*. Retrieved from <https://peerj.com/preprints/3190.pdf>
- [TECH18] Technology, A. I., & Review, I. (2018). *How Random Forest Algorithm Works in Machine Learning What is Random Forest algorithm ? Why Random Forest algorithm ? Random Forest algorithm real life example*.
- [TOWAND] Towards data science. (n.d.). Random Forest Algorithm. Retrieved from <https://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd>

- [VALEND] Valero, M. (n.d.). La otra burbuja en restauración y por qué el capital riesgo se frota las manos. Retrieved from 2018 website:
https://www.elconfidencial.com/empresas/2018-06-06/burbuja-restauracion-capital-riesgo_1573818/
- [VAND18] van der Merwe, R. (2018). *Implementing Facebook Prophet efficiently*. Retrieved from <https://towardsdatascience.com/implementing-facebook-prophet-efficiently-c241305405a3>
- [VINC17] Vincent, T. (2017). A Guide to Time Series Forecasting with Prophet in Python 3. Retrieved from <https://www.digitalocean.com/community/tutorials/a-guide-to-time-series-forecasting-with-prophet-in-python-3>

ANEXOS



8. ANEXOS

ANEXO A – Código Python Random Forest

```
1. #LIBRERIAS
2. import pandas as pd
3. import numpy as np
4. from sklearn.model_selection import train_test_split
5. from sklearn import metrics
6. from sklearn.ensemble import RandomForestRegressor
7. from matplotlib import pyplot
8. import calendar
9.
10. # Import tools needed for visualization
11. from sklearn.tree import export_graphviz
12. import logging
13. logging.getLogger('fbprophet').setLevel(logging.ERROR)
14. import warnings
15. warnings.filterwarnings("ignore")
16. from sklearn.metrics import mean_squared_error
17. from sklearn.model_selection import GridSearchCV
18. from sklearn.metrics import accuracy_score
19. from sklearn.model_selection import cross_val_score
20. pyplot.style.use('fivethirtyeight')
21.
22.
23. #FUNCIONES
24. def cargar_datos(fichero_input):
25.     datos = pd.read_excel(fichero_input)
26.     return datos
27.
28. def generar_fichero_output(resultado, fichero_output):
29.     fichero_prediccion = resultado.to_excel(fichero_output, index = False)
30.     return ()
31.
32.
33. #RANDOM FOREST
34. fichero_input = 'input_historico.xlsx'
35. features = cargar_datos(fichero_input)
36. features.describe()
37.
38. features = pd.get_dummies(features)
39.
40.
41.
42. #RANDOM FOREST
43.
44. #DATAFRAMES
45. labels = features['AFLUENCIA']
46. features = features.drop('AFLUENCIA', axis = 1)
47.
48.
49. #Training data - test data
50. #train_features, test_features, train_labels, test_labels = train_test_split(features
    ,labels,test_size = 0.20, random_state = 42)
51. #TRAIN: 01-01-2018 - 31-12-2018 (365 DIAS)
52. #TEST: 01-01-2019 - 31-01-2019 (31 DIAS)
53.
54.
55. #SEPARAMOS LOS FACTORES EN TRAIN Y TEST
56. cut_date = '2019-02-01'
57. train_features = features[features['ds'] < cut_date]
58. test_features = features[features['ds'] >= cut_date]
59.
```

```
60. features = features.drop('ds', axis = 1)
61. feature_list = list(features.columns)
62.
63. train_features = train_features.drop('ds', axis = 1)
64. test_features = test_features.drop('ds', axis = 1)
65.
66. #SEPARAMOS EL OBJETIVO (AFLUENCIA) EN TRAIN Y TEST
67. train_labels = labels[0:396]
68. test_labels = labels[396:]
69.
70.
71. #CONVERTIMOS A ARRAYS
72. features = np.array(features)
73. labels = np.array(labels)
74.
75. train_features = np.array(train_features)
76. test_features = np.array(test_features)
77. train_labels = np.array(train_labels)
78. test_labels = np.array(test_labels)
79.
80.
81. print('Training Features Shape: ', train_features.shape)
82. print('Training Label Shape: ', train_labels.shape)
83. print('Testing Features Shape: ', test_features.shape)
84. print('Testing Labels Shape: ', test_labels.shape)
85.
86.
87.
88. #ESTABLISH BASELINE
89. # The baseline predictions are the historical averages
90. baseline_preds = test_features[:, feature_list.index('PROMEDIO')]
91.
92. # Baseline errors, and display average baseline error
93. baseline_errors = abs(baseline_preds - test_labels)
94.
95. print('Average baseline error: ', round(np.mean(baseline_errors), 2))
96.
97.
98.
99. #TRAIN MODEL
100.     #Instantiate model with 1000 decision trees
101.     #rf = RandomForestRegressor(n_estimators = 1000, random_state = 42)
102.
103.     rf = RandomForestRegressor(n_estimators = 1000, Random_state = 42)
104.     #Train the model on training data
105.     rf.fit(train_features, train_labels)
106.
107.
108.     #MAKE PREDICTIONS ON THE TEST SET
109.     # Use the forest's predict method on the test data
110.     predictions = rf.predict(test_features)
111.
112.     # Calculate the absolute errors
113.     errors = abs(predictions - test_labels)
114.
115.     #Calculate MAE
116.     mae = 1-(abs(predictions - test_labels)/test_labels)
117.
118.     # Print out the mean absolute error (mae)
119.     print('Mean Absolute Error:', 100-round(np.mean(errors), 2), '%.')
120.     print('MAE:', np.mean(mae)*100)
121.
122.
123.
124.     #DETERMINE PERFORMANCE METRICS
125.     # Calculate mean absolute percentage error (MAPE)
```

```
126.     mape = 100 * (errors / test_labels)
127.
128.     # Calculate and display accuracy
129.     accuracy = 100 - np.mean(mape)
130.     print('Accuracy:', round(accuracy, 2), '%.')
131.
132.
133.
134.     #ANÁLISIS DE RESULTADOS
135.     #INTERPRET MODEL AND REPORT RESULTS
136.     # Get numerical feature importances
137.     importances = list(rf.feature_importances_)
138.
139.     # List of tuples with variable and importance
140.     feature_importances = [(feature, round(importance, 2)) for feature, importance
141. in zip(feature_list, importances)]
141.     # Sort the feature importances by most important first
142.     feature_importances = sorted(feature_importances, key = lambda x: x[1], reverse = True)
143.     # Print out the feature and importances
144.     [print('Variable: {:30} Importance: {}'.format(*pair)) for pair in feature_importances];
145.
146.
147.
148.     #REPRESENTACIÓN DE FIGURAS
149.     #Figura predicción vs real
150.     fig = pyplot.figure()
151.     pyplot.plot(predictions, color = 'red', label = 'Predictions')
152.     pyplot.plot(test_labels, color = 'blue', label = 'Test_labels')
153.
154.     pyplot.legend()
155.     pyplot.title("Predicción afluencia vs Valor real")
156.     pyplot.xlabel('Fecha', fontsize = 10)
157.     pyplot.ylabel('Afluencia comensales', fontsize = 10)
158.
159.     pyplot.show()
160.     #fig.savefig('Forecast_RandomForest_v4.jpg', dpi=1000)
161.
162.
163.     #Figura errores
164.     fig2 = pyplot.figure()
165.     pyplot.plot(errors)
166.     pyplot.title("Errors")
167.     pyplot.show()
168.     #fig2.savefig('Forecast_Error_RandomForest_v3.jpg', dpi=1000)
169.
170.
171.     #ARCHIVO CON DATOS TEST - PREDICCIÓN VS REAL
172.     fechas_test = cargar_datos('input_fechas_test.xlsx')
173.     datos_test = pd.DataFrame(columns = {'ds', 'real', 'prediccion'})
174.     datos_test['ds'] = fechas_test['ds']
175.     for i in range(0,28):
176.         datos_test.loc[i, 'prediccion'] = predictions[i]
177.         datos_test.loc[i, 'real'] = test_labels[i]
178.
179.     datos_test = datos_test[['ds', 'real', 'prediccion']]
180.
181.     fichero_output = 'Datos_Test.xlsx'
182.     generar_fichero_output(datos_test, fichero_output)
183.
184.
185.     #PREDICCIÓN
186.     mes = cargar_datos('input_mes_prediccion.xlsx')
187.     prediccion_mes = rf.predict(mes)
188.
```

```
189.     cols = { 'AFLUENCIA PRED' }
190.     d_prediccion_mes = pd.DataFrame(prediccion_mes, columns= cols)
191.
192.     d_prediccion_mes['DIA']= mes['DIA']
193.     d_prediccion_mes['MES']= mes['MES']
194.     d_prediccion_mes['AÑO'] = mes['AÑO']
195.
196.     d_prediccion_mes[['DIA', 'MES', 'AÑO', 'AFLUENCIA PRED']]
197.
198.     fichero_output = 'Output_Marzo_RandomForest.xlsx'
199.     #generar_fichero_output(d_prediccion_mes, fichero_output)
200.
201.     d_prediccion_mes['AFLUENCIA PRED'].plot()
202.
203.
204.     generar_fichero_output(d_prediccion_mes, fichero_output)
```

ANEXO B – Código Python Prophet

```
1. #LIBRERÍA
2. import pandas as pd
3. import numpy as np
4. from fbprophet import Prophet
5. import matplotlib.pyplot as plt
6. plt.style.use('fivethirtyeight')
7. from fbprophet.diagnostics import performance_metrics
8. from fbprophet.diagnostics import cross_validation
9. from fbprophet.plot import plot_cross_validation_metric
10. from fbprophet.plot import plot_yearly
11. from sklearn.model_selection import cross_val_score
12. import logging
13. logging.getLogger('fbprophet').setLevel(logging.ERROR)
14. import warnings
15. warnings.filterwarnings("ignore")
16. from sklearn.metrics import mean_squared_error
17. from sklearn.model_selection import GridSearchCV
18. from sklearn.metrics import accuracy_score
19.
20.
21. #FUNCIONES
22. def cargar_datos(fichero_input):
23.     datos = pd.read_excel(fichero_input)
24.     #datos = read.csv(fichero_input, 'r', 'utf-8')
25.
26.     return datos
27.
28.
29. def generar_fichero_output(resultado, fichero_output):
30.     fichero_prediccion = resultado.to_excel(fichero_output, index = False)
31.
32.     return ()
33.
34. def make_comparison_dataframe(historical, forecast):
35.     return forecast.set_index('ds')[['yhat', 'yhat_lower', 'yhat_upper']].join(historical.set_index('ds'))
36.
37.
38. def calculate_forecast_errors(df, prediction_size):
39.     df = df.copy()
40.
41.     df['e'] = df['y'] - df['yhat']
42.     df['p'] = 100 * df['e'] / df['y']
43.
44.     predicted_part = df[-prediction_size:]
45.
46.     error_mean = lambda error_name: np.mean(np.abs(predicted_part[error_name]))
47.
48.     return error_mean('p'), error_mean('e')
49.
50. def calculate_MASE(value_set, prediction_set):
51.
52.     e = np.abs(value_set - prediction_set)
53.     n = value_set.shape[0]
54.     d = np.diff(value_set).sum() / (n-1)
55.
56.     error = np.abs(e / d).mean()
57.
58.     return error
59.
60.
61.
62. #PROPHET
63. fichero_input = 'input_historico_restaurante1_afluencia_v4.xlsx'
```

```

64. df = cargar_datos(fichero_input)
65.
66. #Train set y Test set
67. cut_date = '2019-02-01'
68. df_1 = df[df['ds'] < cut_date]
69. df_2 = df[df['ds'] >= cut_date]
70.
71.
72. #Input factores
73. holidays = pd.DataFrame({'holiday': 'Holiday', 'ds': pd.to_datetime(list(pd.read_csv(
    'Festivos.csv')['ds'])), format = '%d/%m/%Y'), 'lower_window': 0, 'upper_window': 1,})
74. lluvias = pd.DataFrame({'holiday': 'Raining', 'ds': pd.to_datetime(list(pd.read_csv('
    Lluvias.csv')['ds'])), format = '%d/%m/%Y'), 'lower_window': 0, 'upper_window': 1,})
75. temp = pd.DataFrame({'holiday': 'TemperaturaAlta', 'ds': pd.to_datetime(list(pd.read_
    csv('Temperaturas_altas.csv')['ds'])), format = '%d/%m/%Y'), 'lower_window': 0, 'upper_
    window': 0,})
76.
77. all_events = pd.concat((holidays, lluvias, temp ))
78. all_events = all_events.sort_values(['ds'])
79.
80.
81. #Modelo
82. m = Prophet(yearly_seasonality = True, weekly_seasonality = True, seasonality_prior_s
    cale=0.1, changepoint_prior_scale = 0.5, holidays = all_events, holidays_prior_scale
    = 0.1, interval_width = 0.95)
83. m.add_seasonality(name='monthly', period=30.5, fourier_order=5)
84. m.fit(df_1)
85.
86.
87. future= m.make_future_dataframe(periods = 28, freq = 'D', include_history = True)
88.
89.
90. forecast = m.predict(future)
91. forecast_1 = forecast[forecast['ds'] < cut_date]
92. forecast_2 = forecast[forecast['ds'] >= cut_date]
93.
94.
95.
96.
97. #ANÁLISIS RESULTADOS
98. #Visualización resultados test - predicción y valor real
99. d = pd.concat([forecast_2['yhat'], df_2['y']], axis = 1)
100.     d
101.
102.     #Tabla errores modelo
103.     dt = {}
104.     cols = {'MSE_train', 'MAE_train', 'MAPE_train', 'MSE_test', 'MAE_test', 'MAPE_
        test', 'MASE_train', 'MASE_test'}
105.     analysis = pd.DataFrame(columns = cols)
106.
107.     MSE_train = mean_squared_error(df_1['y'], forecast_1['yhat'])
108.     MSE_test = mean_squared_error(df_2['y'], forecast_2['yhat'])
109.
110.     cmp_df1 = make_comparison_dataframe(df_1, forecast_1)
111.     cmp_df2 = make_comparison_dataframe(df_2, forecast_2)
112.
113.     MAE_train, MAPE_train = calculate_forecast_errors(cmp_df1, 316)
114.     MAE_test, MAPE_test = calculate_forecast_errors(cmp_df2, 80)
115.
116.     MASE_train = calculate_MASE(df_1['y'], forecast_1['yhat'])
117.     MASE_test = calculate_MASE(df_2['y'], forecast_2['yhat'])
118.
119.     dt['MSE'] = {'train': MSE_train, 'test': MSE_test }
120.     dt['MAE'] = {'train': MAE_train, 'test': MAE_test }
121.     dt['MAPE'] = {'train': MAPE_train, 'test': MAPE_test }

```

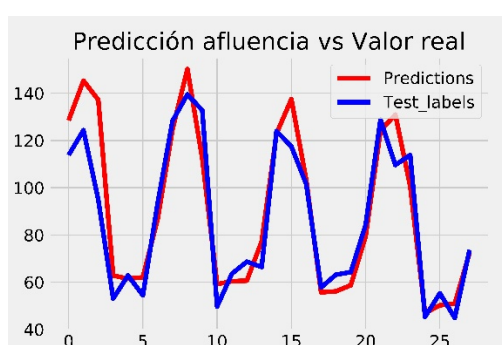
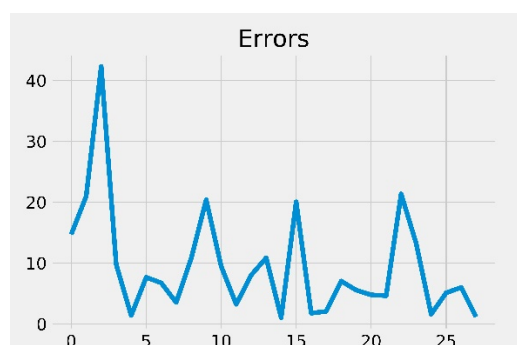
```
122.     dt['MASE'] = {'train': MASE_train, 'test': MASE_test }
123.
124.     dt['MSE_train'] = MSE_train
125.     dt['MAE_train'] = MAE_train
126.     dt['MAPE_train'] = MAPE_train
127.     dt['MSE_test'] = MSE_test
128.     dt['MAE_test'] = MAE_test
129.     dt['MAPE_test'] = MAPE_test
130.     dt['MASE_train'] = MASE_train
131.     dt['MASE_test'] = MASE_test
132.
133.     analysis.loc[0, cols] = dt
134.
135.
136.     analysis[['MAPE_train', 'MAE_train', 'MSE_train', 'MASE_train', 'MAPE_test', 'MAE
_test', 'MSE_test', 'MASE_test']]
137.
138.
139.     #Exportamos datos del test para validar el accuracy
140.     test_datos = cmp_df2[['yhat', 'y']]
141.     fichero_output = 'Datos_Test_v7.xlsx'
142.     generar_fichero_output(test_datos, fichero_output)
143.
144.
145.
146.     #Representación histórico
147.     ax = df.set_index('ds').plot(figsize=(12,8))
148.     ax.set_ylabel('Numero de comensales')
149.     ax.set_xlabel('Fecha')
150.
151.     plt.show()
152.
153.
154.     #Representación forecast
155.     figure = m.plot(forecast)
156.     for changepoint in m.changepoints:
157.         plt.axvline(changepoint, ls='--', lw=1)
158.
159.     #figure.savefig('Forecast_Prophet_v5.jpg', dpi=1000)
160.
161.     #Representación componentes
162.     m.plot(forecast, uncertainty=True)
163.     m.plot_components(forecast)
164.
165.
166.     #FORECAST
167.     marzo = cargar_datos('input_calendario_marzo.xlsx')
168.     forecast_marzo = m.predict(marzo)
169.     forecast_marzo[['ds', 'yhat']]
170.
171.     fichero_output = 'Output_Marzo_Prophet_v1.xlsx'
172.     generar_fichero_output(forecast_marzo, fichero_output)
173.
174.
175.     #PREDICCIÓN FEBRERO
176.     forecast_febrero['yhat'].plot()
```


ANEXO C – Importancia de variables según Random Forest

RESTAURANTE 1

Training Features Shape: (396, 20)
 Training Label Shape: (396,)
 Testing Features Shape: (28, 20)
 Testing Labels Shape: (28,)
 Average baseline error: 29.39
 Mean Absolute Error: 90.53 %.
 MAE: 89.10279199889014
 Accuracy: 89.1 %.

Variable: DIA_SEMANA_sábado	Importance: 0.28
Variable: DIA_SEMANA_viernes	Importance: 0.25
Variable: DIA_SEMANA_domingo	Importance: 0.19
Variable: TMED	Importance: 0.13
Variable: DIA	Importance: 0.04
Variable: MES	Importance: 0.03
Variable: PRECIPITACION	Importance: 0.03
Variable: DIA_SEMANA_jueves	Importance: 0.03
Variable: FESTIVO	Importance: 0.02
Variable: AÑO	Importance: 0.0
Variable: FESTIVO_NAVIDADES	Importance: 0.0
Variable: CAPACIDAD	Importance: 0.0
Variable: PROMEDIO	Importance: 0.0
Variable: DIA_SEMANA_lunes	Importance: 0.0
Variable: DIA_SEMANA_martes	Importance: 0.0
Variable: DIA_SEMANA_miércoles	Importance: 0.0
Variable: RESTAURANTE_Restaurante 1	Importance: 0.0
Variable: TIPO_CC	Importance: 0.0
Variable: PARKING_si	Importance: 0.0
Variable: TERRAZA_no	Importance: 0.0

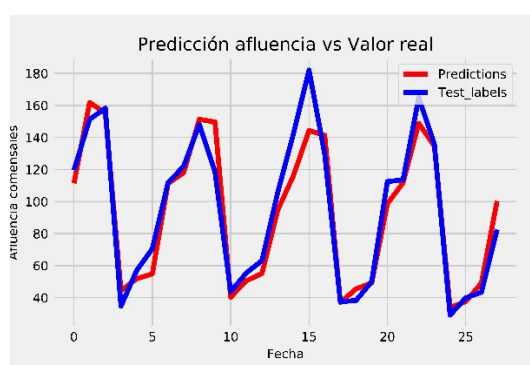


RESTAURANTE 2

Training Features Shape: (396, 20)
 Training Label Shape: (396,)
 Testing Features Shape: (28, 20)
 Testing Labels Shape: (28,)
 Average baseline error: 41.98
 Mean Absolute Error: 90.51 %.
 MAE: 89.36502398223531
 Accuracy: 89.37 %.

Variable: DIA_SEMANA_sábado	Importance: 0.24
Variable: DIA_SEMANA_domingo	Importance: 0.17
Variable: DIA_SEMANA_viernes	Importance: 0.14

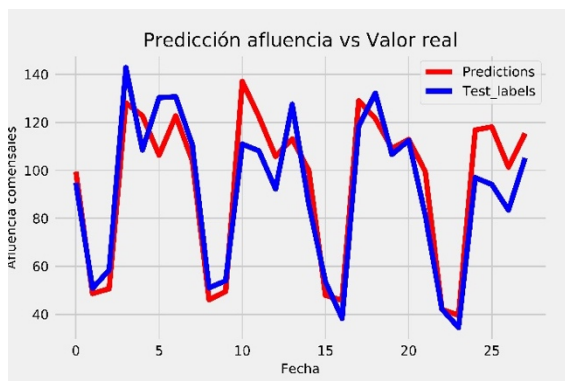
Variable: DIA_SEMANA_jueves	Importance: 0.12
Variable: TMED	Importance: 0.07
Variable: DIA_SEMANA_lunes	Importance: 0.05
Variable: DIA_SEMANA_miércoles	Importance: 0.05
Variable: DIA	Importance: 0.04
Variable: MES	Importance: 0.04
Variable: DIA_SEMANA_martes	Importance: 0.04
Variable: PRECIPITACION	Importance: 0.03
Variable: FESTIVO	Importance: 0.01
Variable: AÑO	Importance: 0.0
Variable: FESTIVO_NAVIDADES	Importance: 0.0
Variable: CAPACIDAD	Importance: 0.0
Variable: PROMEDIO	Importance: 0.0
Variable: RESTAURANTE_Restaurante 2	Importance: 0.0
Variable: TIPO_R	Importance: 0.0
Variable: PARKING_no	Importance: 0.0
Variable: TERRAZA_si	Importance: 0.0



RESTAURANTE 3

Training Features Shape: (396, 20)
 Training Label Shape: (396,)
 Testing Features Shape: (28, 20)
 Testing Labels Shape: (28,)
 Average baseline error: 26.01
 Mean Absolute Error: 88.99 %.
 MAE: 87.81701998088496
 Accuracy: 87.82 %.

Variable: DIA_SEMANA_domingo	Importance: 0.33
Variable: DIA_SEMANA_sábado	Importance: 0.33
Variable: TMED	Importance: 0.09
Variable: DIA	Importance: 0.06
Variable: FESTIVO	Importance: 0.06
Variable: MES	Importance: 0.05
Variable: DIA_SEMANA_viernes	Importance: 0.03
Variable: FESTIVO_NAVIDADES	Importance: 0.01
Variable: PRECIPITACION	Importance: 0.01
Variable: DIA_SEMANA_lunes	Importance: 0.01
Variable: DIA_SEMANA_miércoles	Importance: 0.01
Variable: AÑO	Importance: 0.0
Variable: CAPACIDAD	Importance: 0.0
Variable: PROMEDIO	Importance: 0.0
Variable: DIA_SEMANA_jueves	Importance: 0.0
Variable: DIA_SEMANA_martes	Importance: 0.0
Variable: RESTAURANTE_Restaurante 3	Importance: 0.0
Variable: TIPO_O	Importance: 0.0
Variable: PARKING_no	Importance: 0.0
Variable: TERRAZA_no	Importance: 0.0

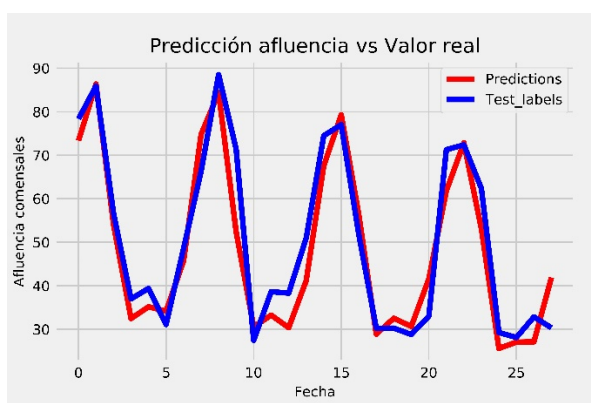


RESTAURANTE 4

Training Features Shape: (396, 20)
 Training Label Shape: (396,)
 Testing Features Shape: (28, 20)
 Testing Labels Shape: (28,)
 Average baseline error: 18.24
 Mean Absolute Error: 94.66 %.
 MAE: 88.3748365566982
 Accuracy: 88.37 %.

Variable: DIA_SEMANA_sábado
 Variable: DIA_SEMANA_viernes
 Variable: DIA_SEMANA_domingo
 Variable: TMED
 Variable: DIA
 Variable: PRECIPITACION
 Variable: DIA_SEMANA_jueves
 Variable: MES
 Variable: FESTIVO
 Variable: FESTIVO_NAVIDADES
 Variable: AÑO
 Variable: CAPACIDAD
 Variable: PROMEDIO
 Variable: DIA_SEMANA_lunes
 Variable: DIA_SEMANA_martes
 Variable: DIA_SEMANA_miércoles
 Variable: RESTAURANTE_Restaurante 4
 Variable: TIPO_CC
 Variable: PARKING_si
 Variable: TERRAZA_si

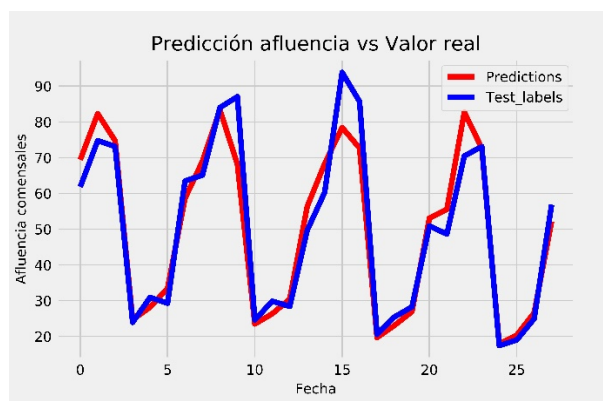
Importance: 0.3
 Importance: 0.27
 Importance: 0.16
 Importance: 0.08
 Importance: 0.07
 Importance: 0.04
 Importance: 0.04
 Importance: 0.03
 Importance: 0.01
 Importance: 0.01
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0
 Importance: 0.0



RESTAURANTE 5

Training Features Shape: (396, 20)
 Training Label Shape: (396,)
 Testing Features Shape: (28, 20)
 Testing Labels Shape: (28,)
 Average baseline error: 21.49
 Mean Absolute Error: 95.1 %.
 MAE: 91.11709548062181
 Accuracy: 91.12 %.

Variable: DIA_SEMANA_sábado	Importance: 0.23
Variable: DIA_SEMANA_domingo	Importance: 0.19
Variable: DIA_SEMANA_viernes	Importance: 0.14
Variable: DIA_SEMANA_jueves	Importance: 0.12
Variable: TMED	Importance: 0.11
Variable: DIA_SEMANA_lunes	Importance: 0.04
Variable: DIA_SEMANA_miércoles	Importance: 0.04
Variable: DIA	Importance: 0.03
Variable: MES	Importance: 0.03
Variable: PRECIPITACION	Importance: 0.03
Variable: DIA_SEMANA_martes	Importance: 0.03
Variable: FESTIVO	Importance: 0.01
Variable: AÑO	Importance: 0.0
Variable: FESTIVO_NAVIDADES	Importance: 0.0
Variable: CAPACIDAD	Importance: 0.0
Variable: PROMEDIO	Importance: 0.0
Variable: RESTAURANTE_Restaurante 5	Importance: 0.0
Variable: TIPO_R	Importance: 0.0
Variable: PARKING_si	Importance: 0.0
Variable: TERRAZA_no	Importance: 0.0

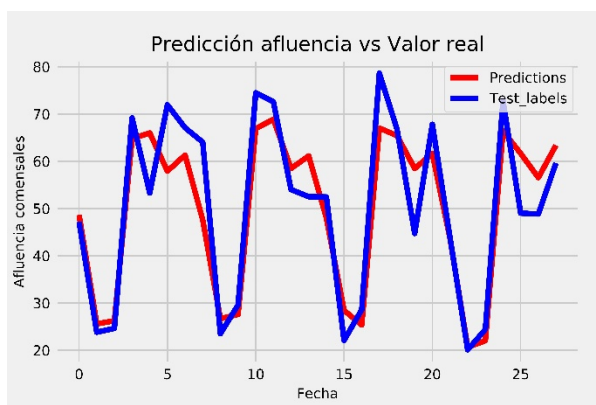


RESTAURANTE 6

Training Features Shape: (396, 20)
 Training Label Shape: (396,)
 Testing Features Shape: (28, 20)
 Testing Labels Shape: (28,)
 Average baseline error: 15.82
 Mean Absolute Error: 93.99 %.
 MAE: 87.99027219376642
 Accuracy: 87.99 %.

Variable: DIA_SEMANA_sábado	Importance: 0.33
Variable: DIA_SEMANA_domingo	Importance: 0.32
Variable: TMED	Importance: 0.08

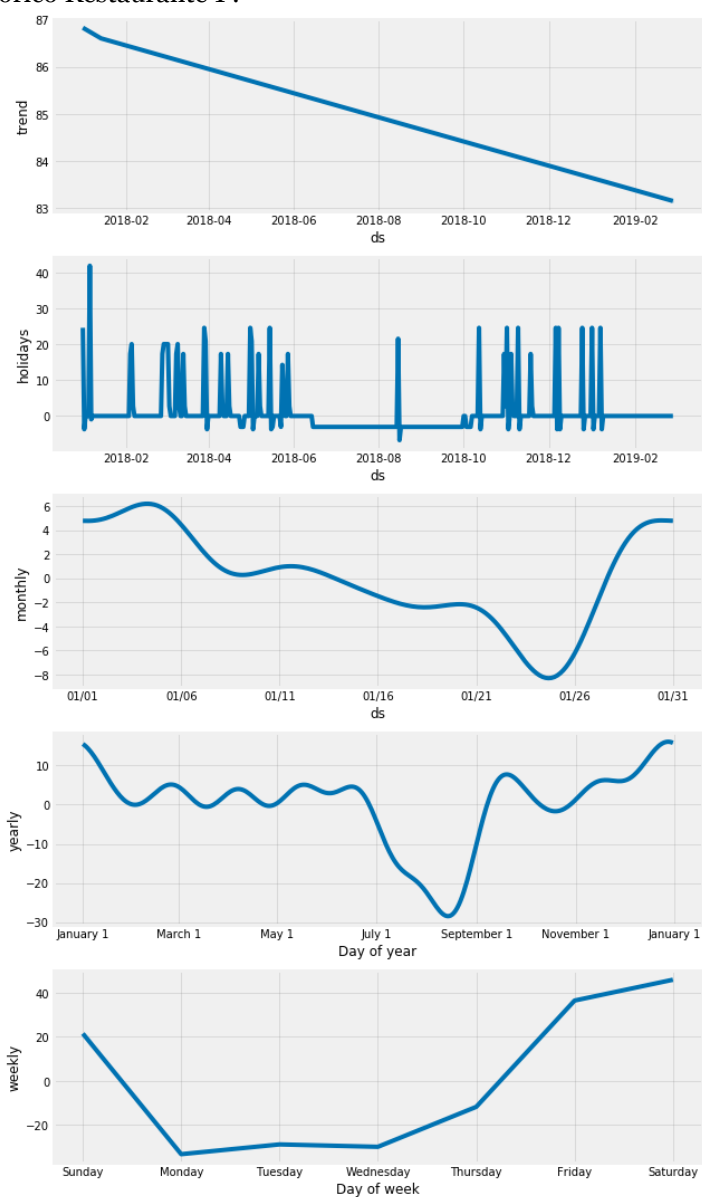
Variable: MES	Importance: 0.06
Variable: FESTIVO	Importance: 0.06
Variable: DIA	Importance: 0.05
Variable: DIA_SEMANA_viernes	Importance: 0.05
Variable: FESTIVO_NAVIDADES	Importance: 0.01
Variable: PRECIPITACION	Importance: 0.01
Variable: DIA_SEMANA_lunes	Importance: 0.01
Variable: DIA_SEMANA_martes	Importance: 0.01
Variable: DIA_SEMANA_miércoles	Importance: 0.01
Variable: AÑO	Importance: 0.0
Variable: CAPACIDAD	Importance: 0.0
Variable: PROMEDIO	Importance: 0.0
Variable: DIA_SEMANA_jueves	Importance: 0.0
Variable: RESTAURANTE_Restaurante 6	Importance: 0.0
Variable: TIPO_O	Importance: 0.0
Variable: PARKING_no	Importance: 0.0
Variable: TERRAZA_si	Importance: 0.0



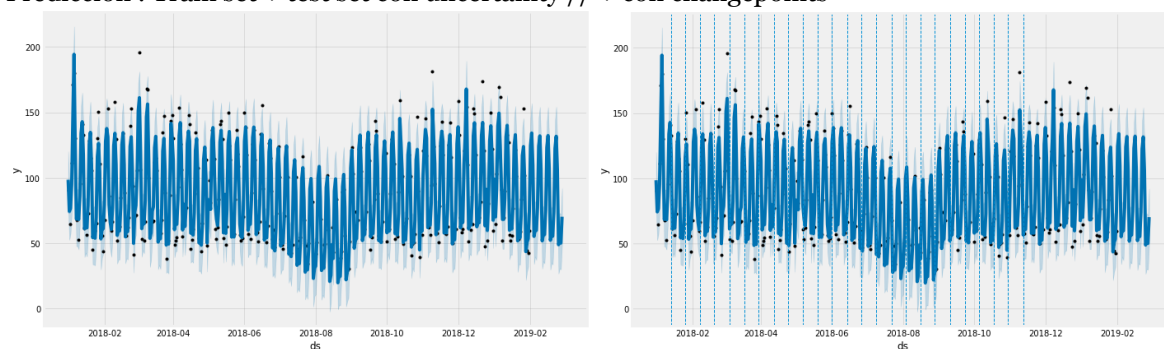
ANEXO D – Gráficas obtenidas del Modelo Prophet

RESTAURANTE 1

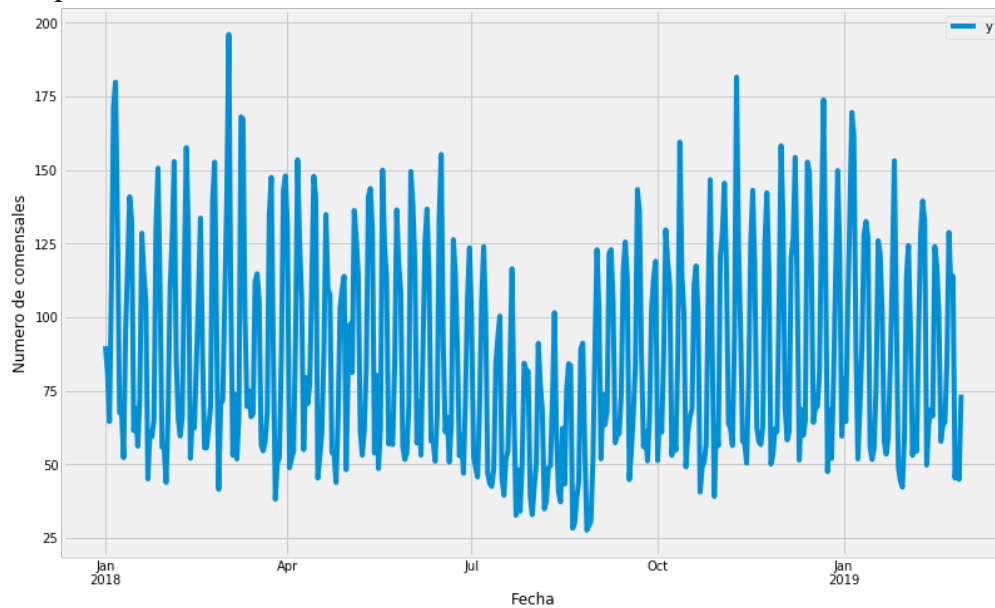
Componentes del histórico Restaurante 1 :



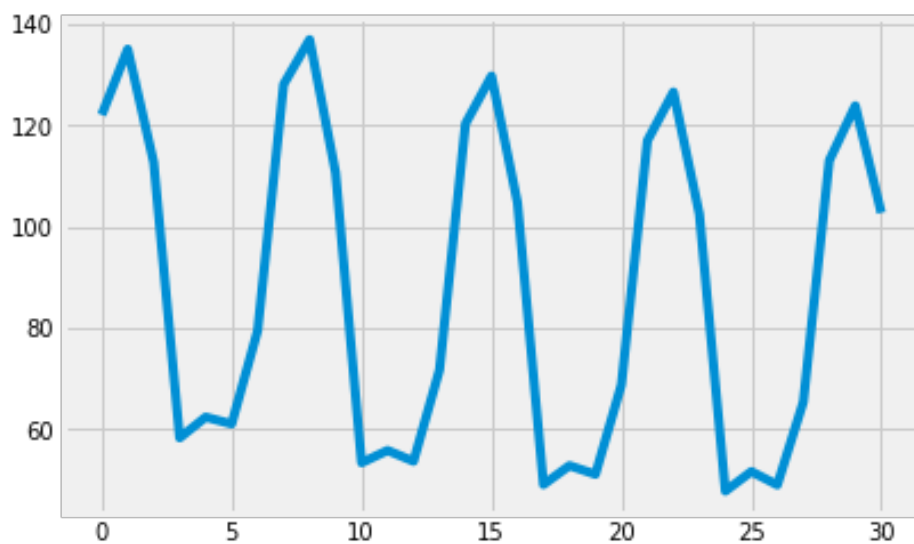
Predicción : Train set + test set con uncertainty // + con changepoints



Histórico sin predecir:

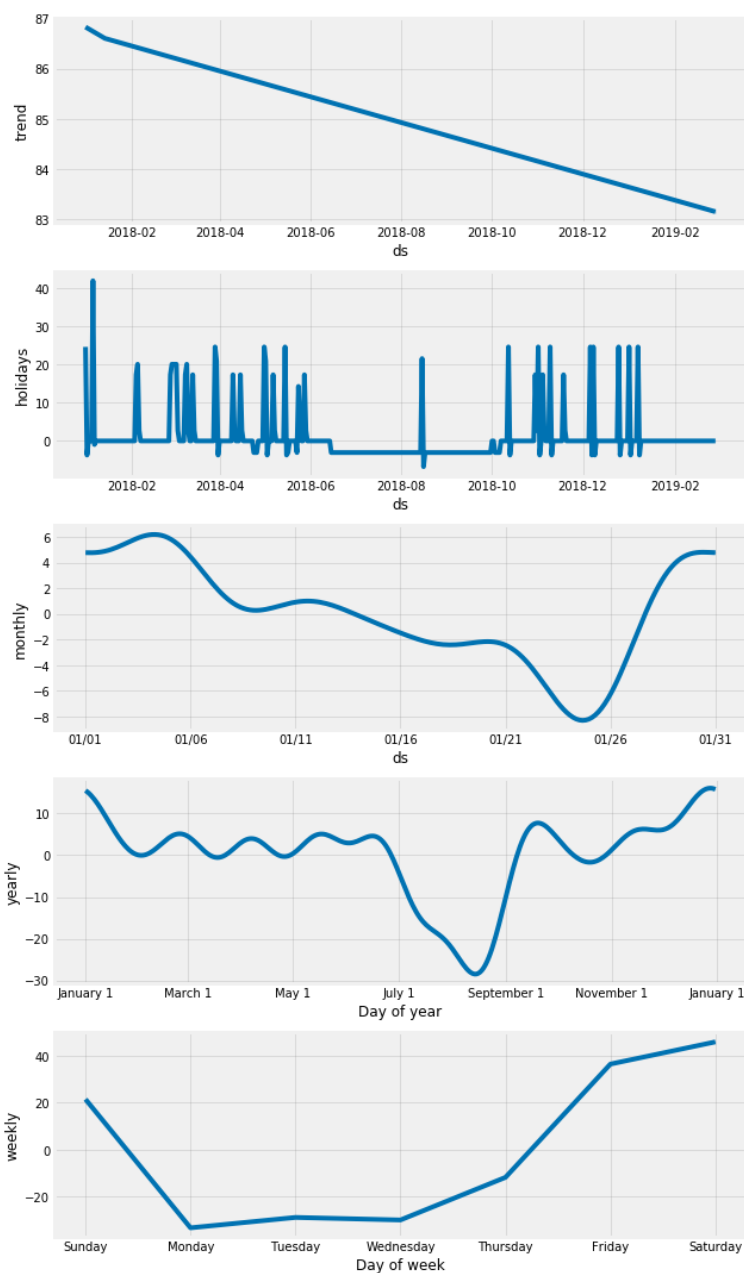


Predicción Marzo:

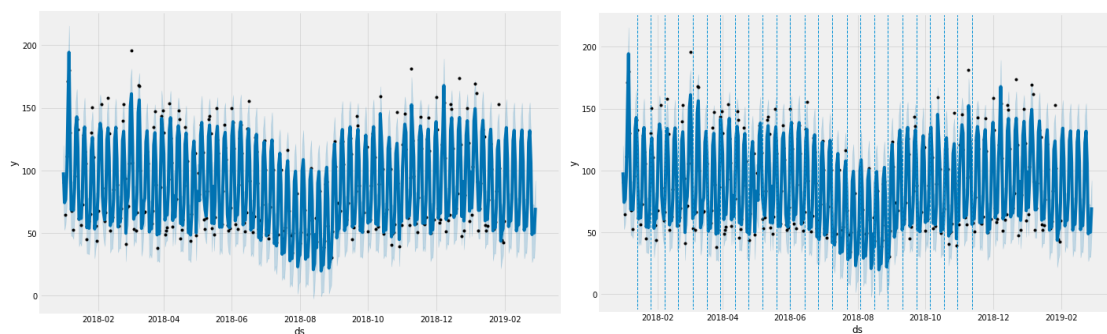


RESTAURANTE 2

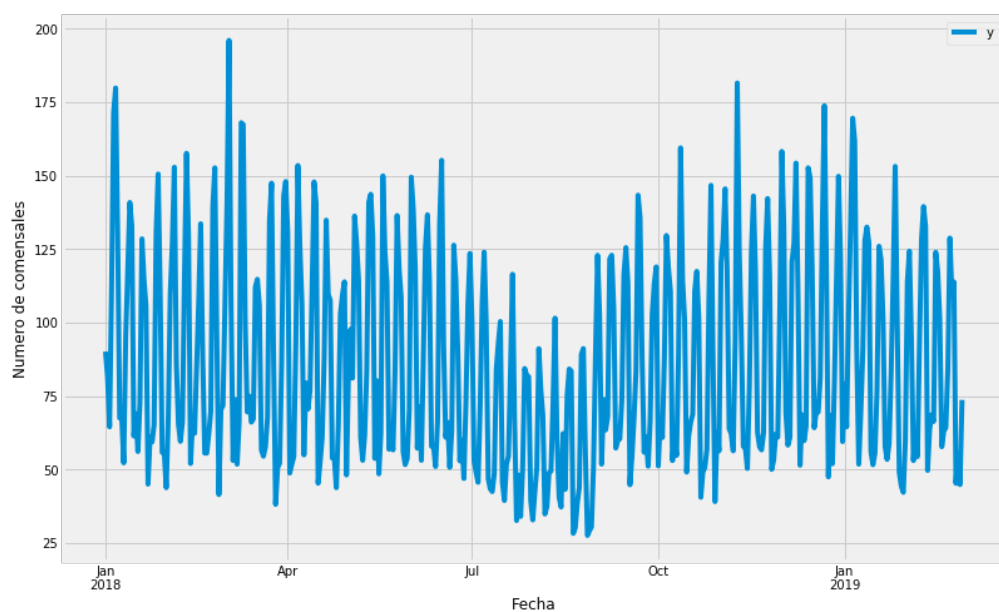
Componentes del histórico Restaurante 2:



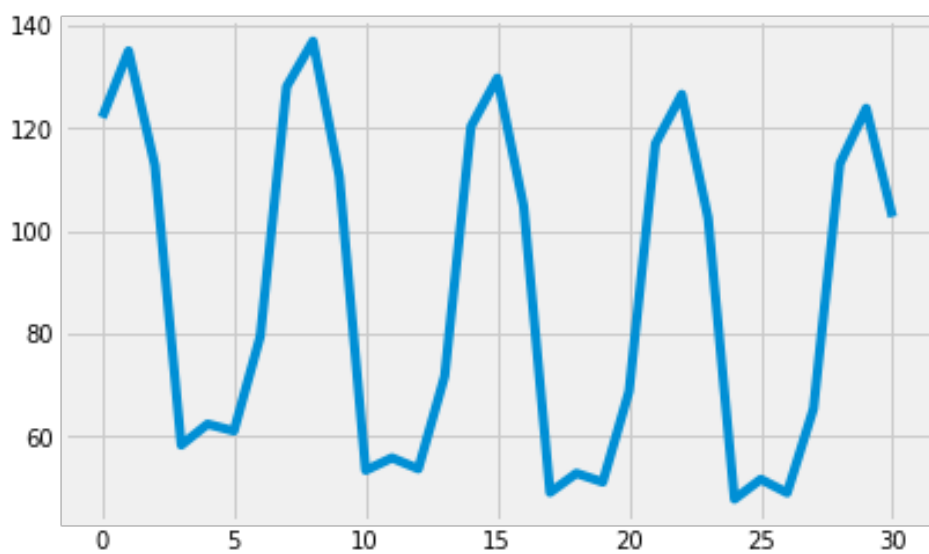
Predicción : Train set + test set con uncertainty // + con changepoints



Histórico sin predecir:

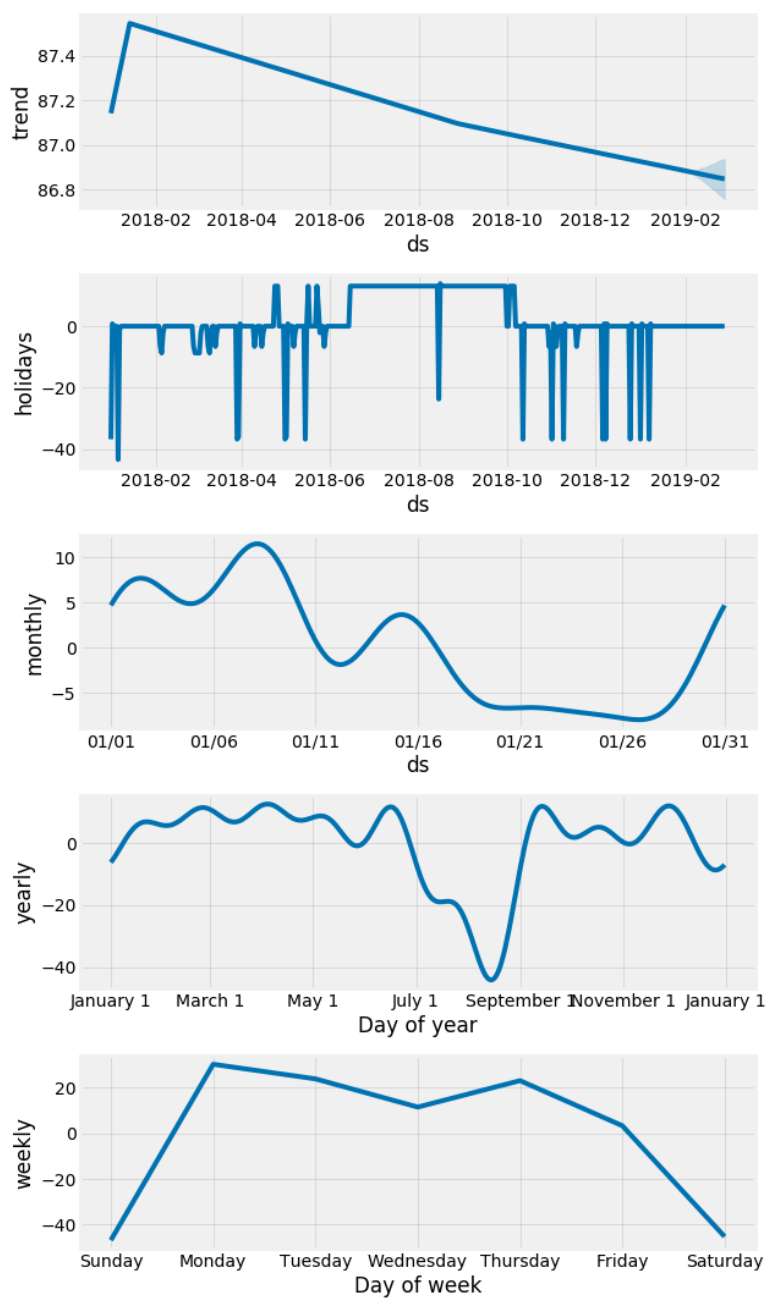


Predicción Marzo:

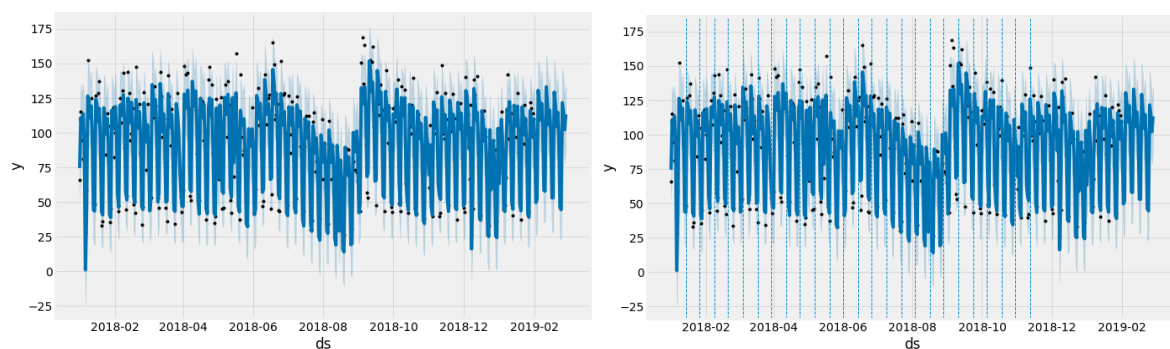


RESTAURANTE 3

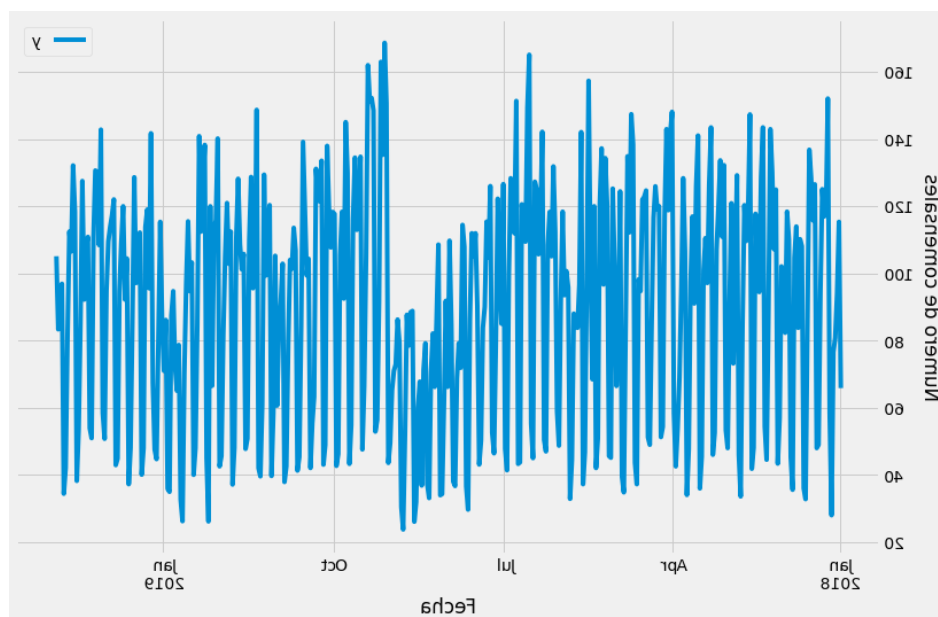
Componentes del histórico Restaurante 3



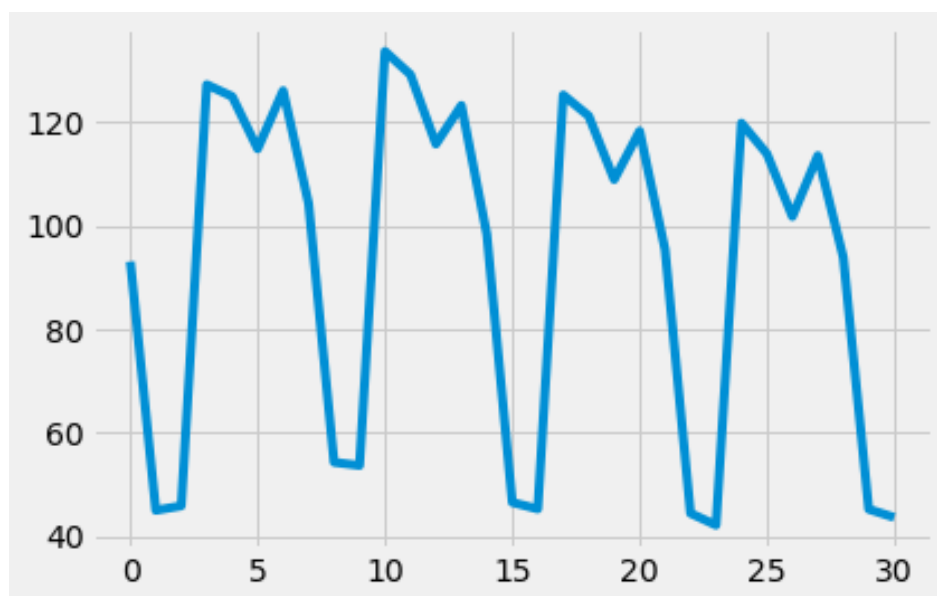
Predicción : Train set + test set con uncertainty // + con changepoints



Histórico sin predecir:

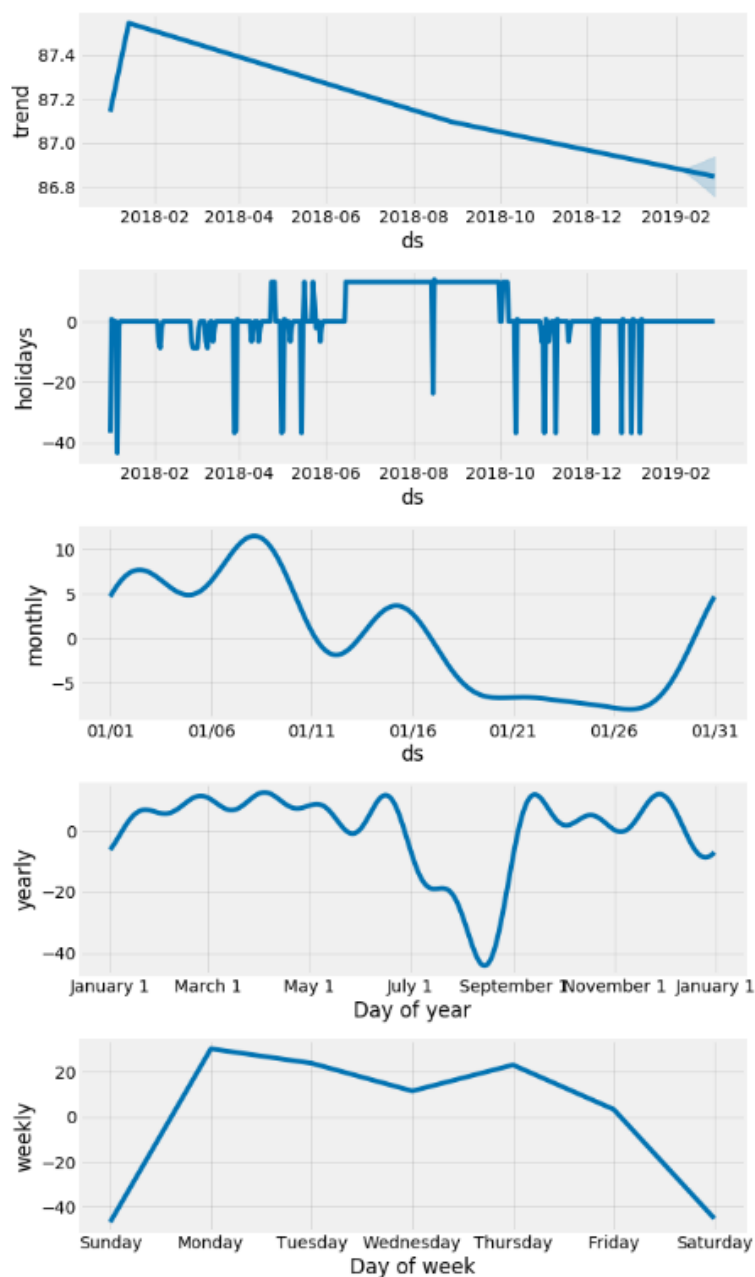


Predicción Marzo:

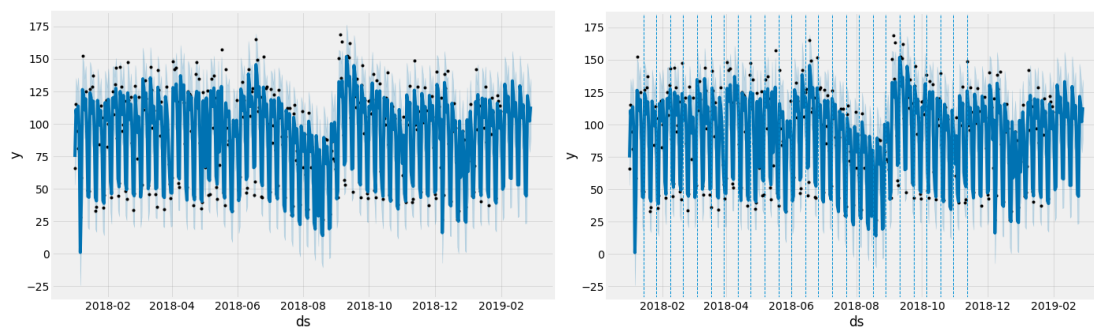


RESTAURANTE 4

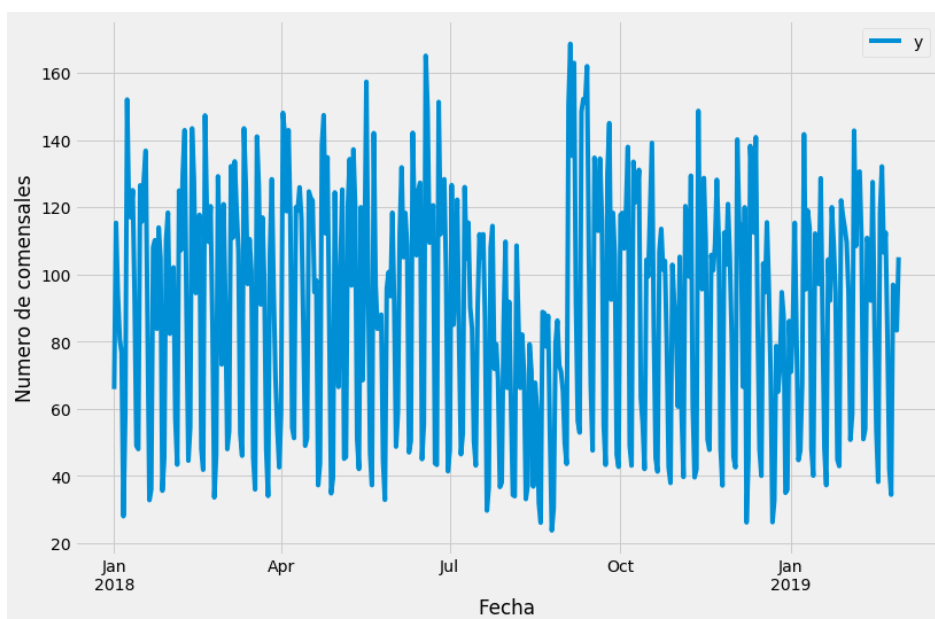
Componentes del histórico Restaurante 4



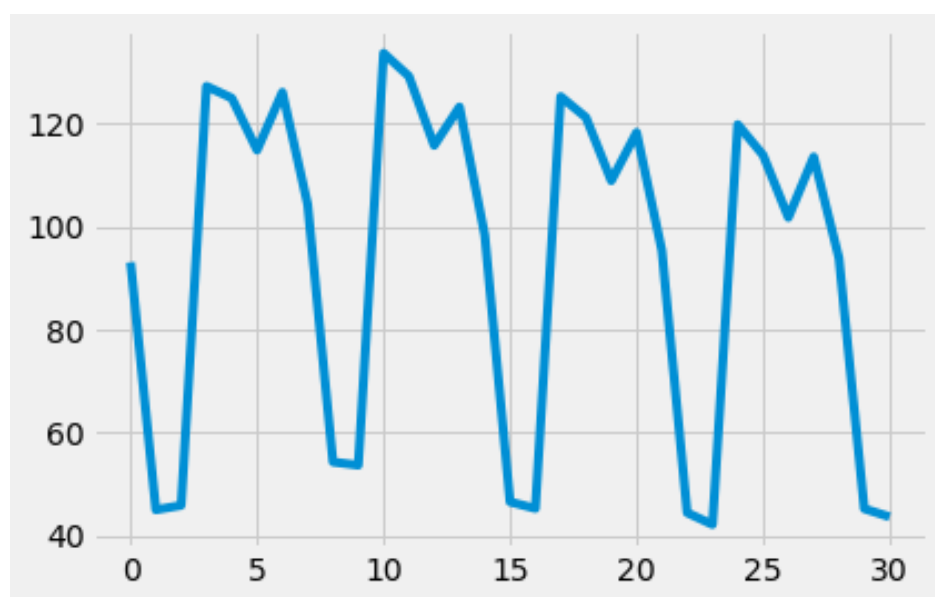
Predicción : Train set + test set con uncertainty // + con changepoints



Histórico sin predecir:

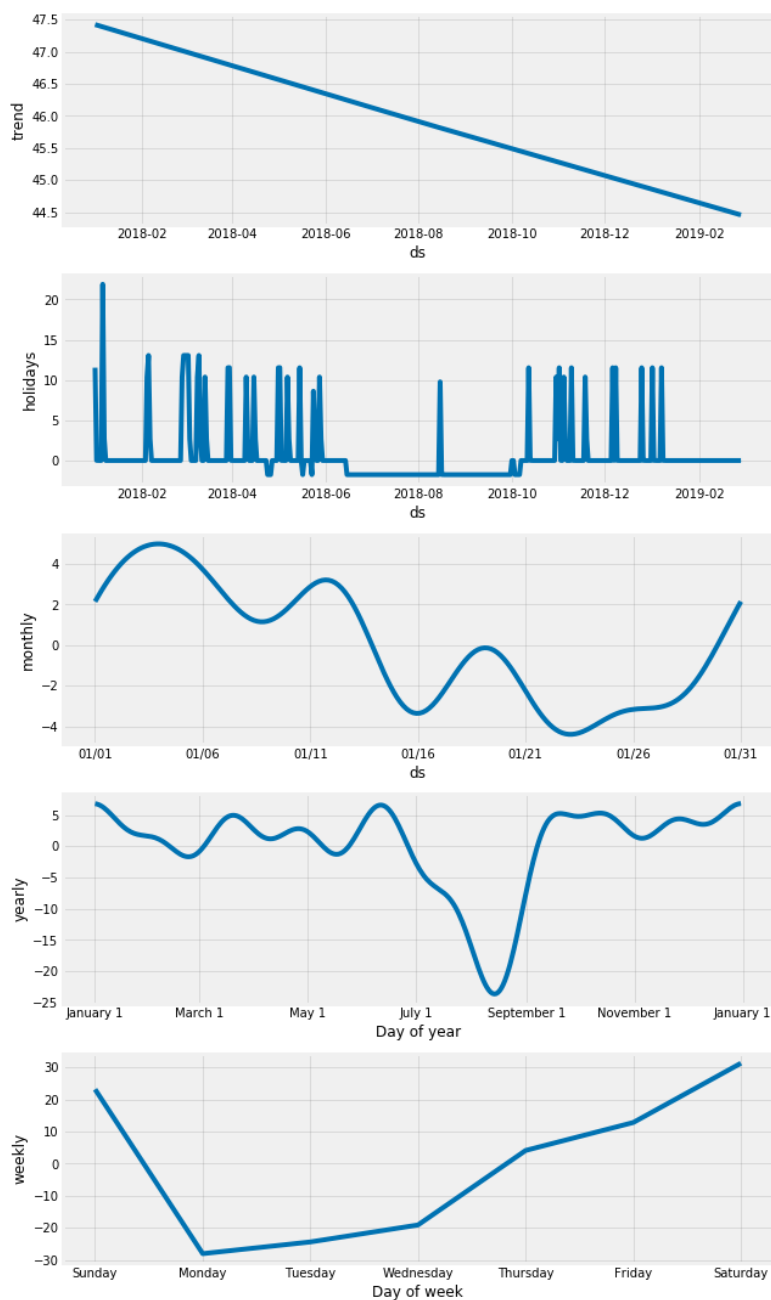


Predicción Marzo:

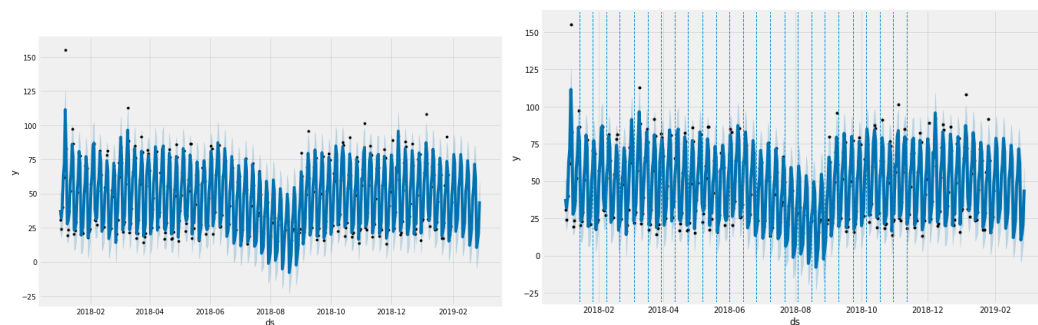


RESTAURANTE 5:

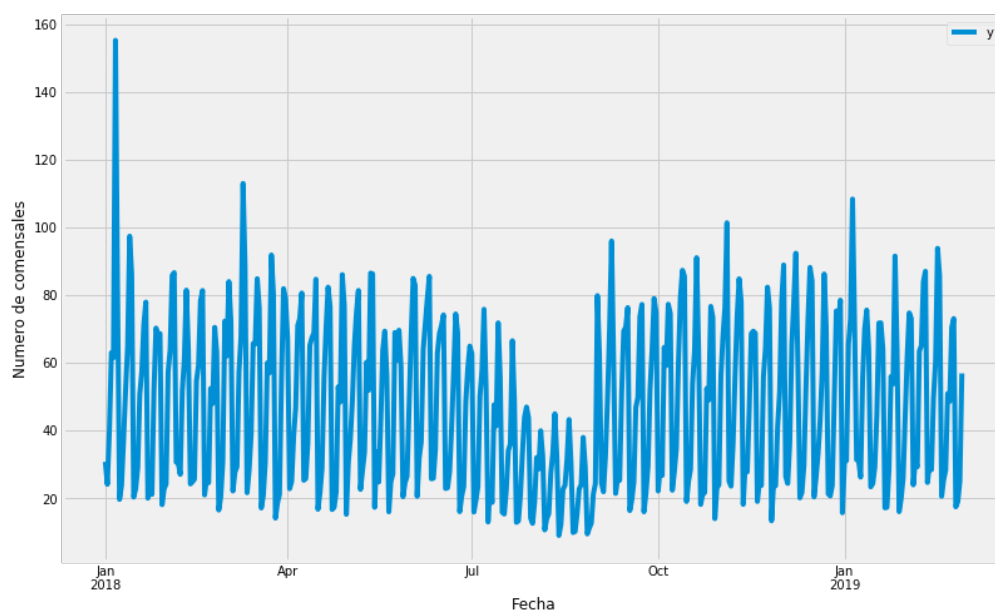
Componentes del histórico Restaurante 5:



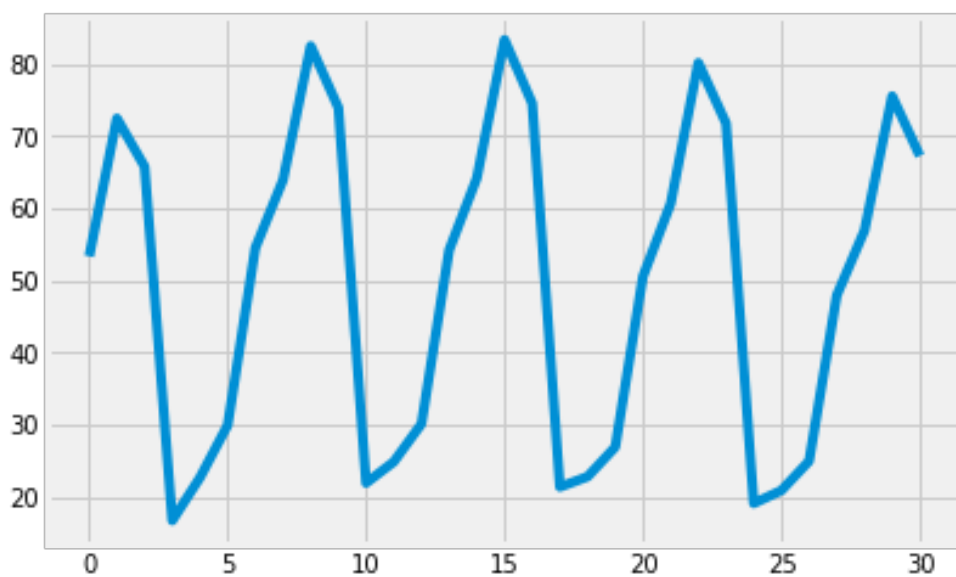
Predicción : Train set + test set con uncertainty // + con changepoints



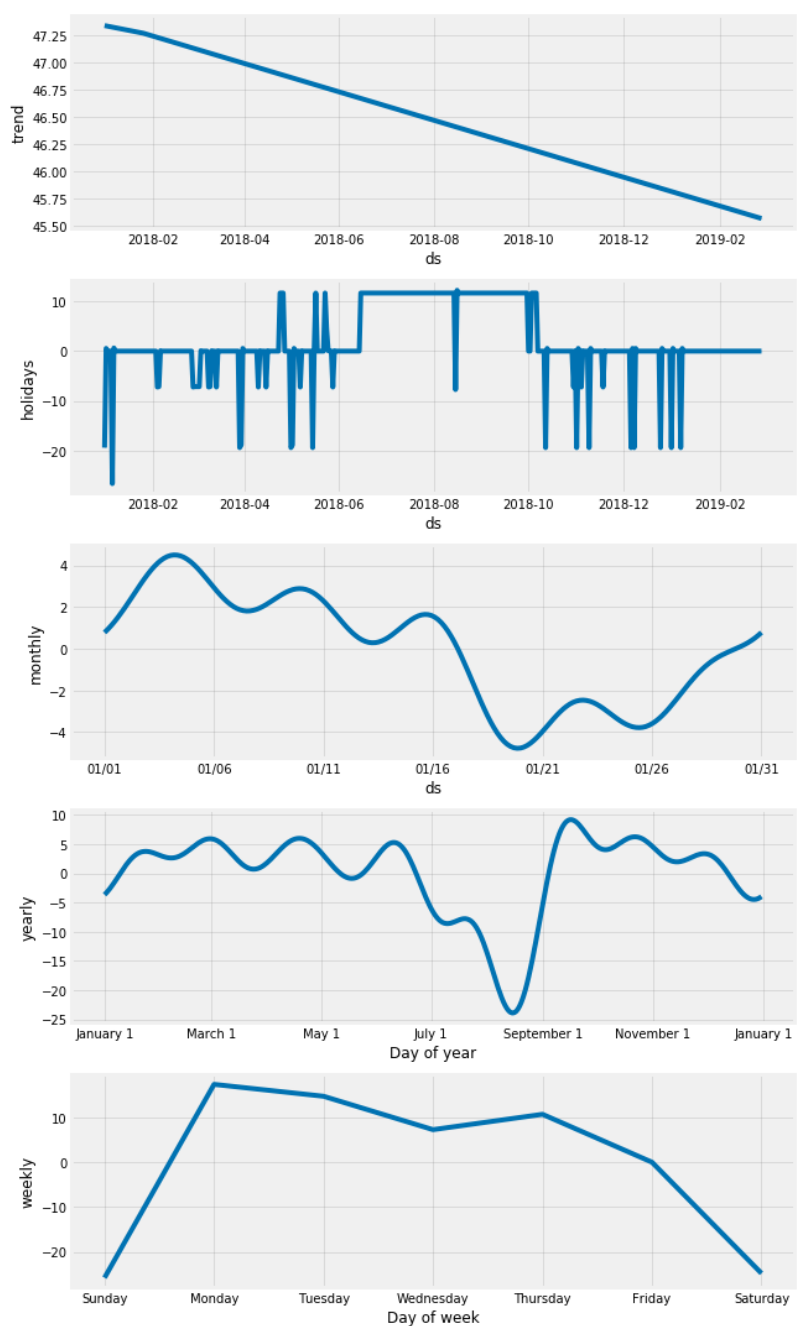
Histórico sin predecir:



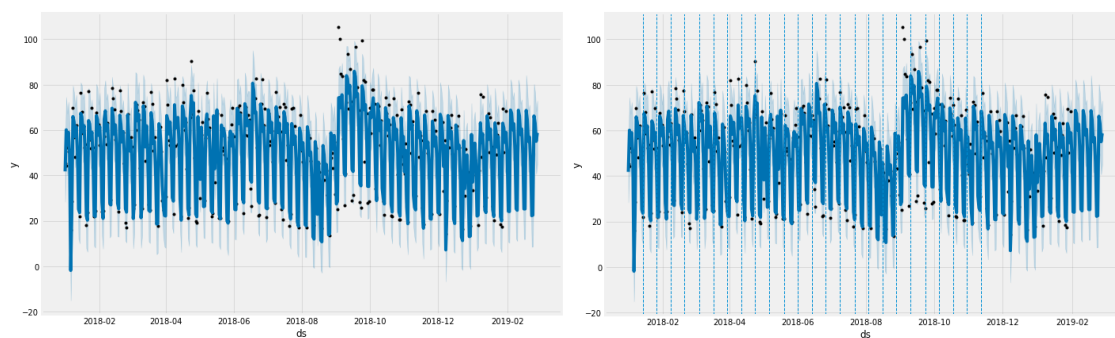
Predicción Marzo:



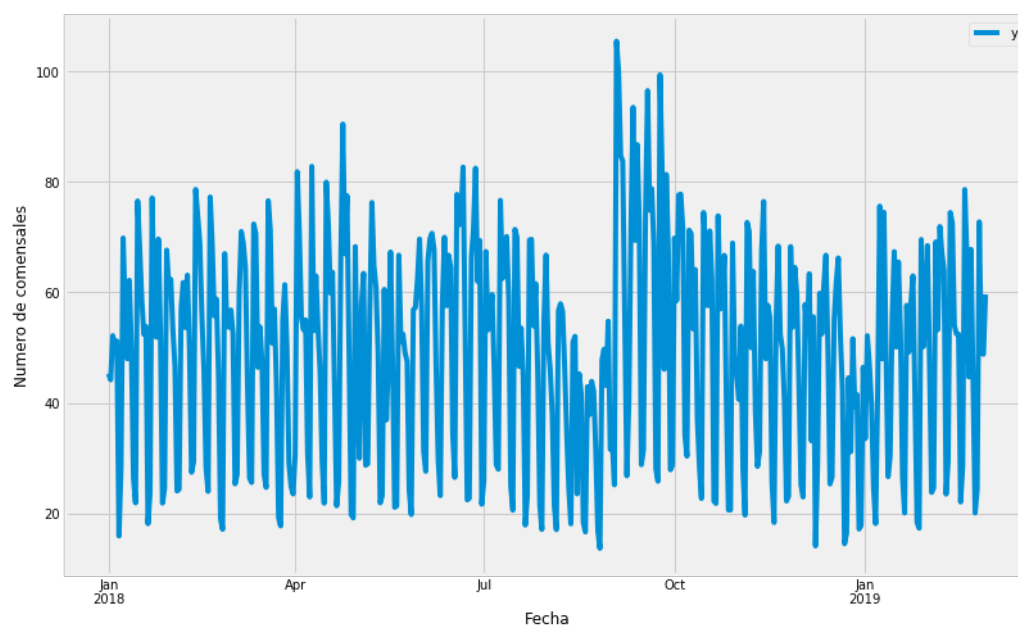
RESTAURANTE 6:



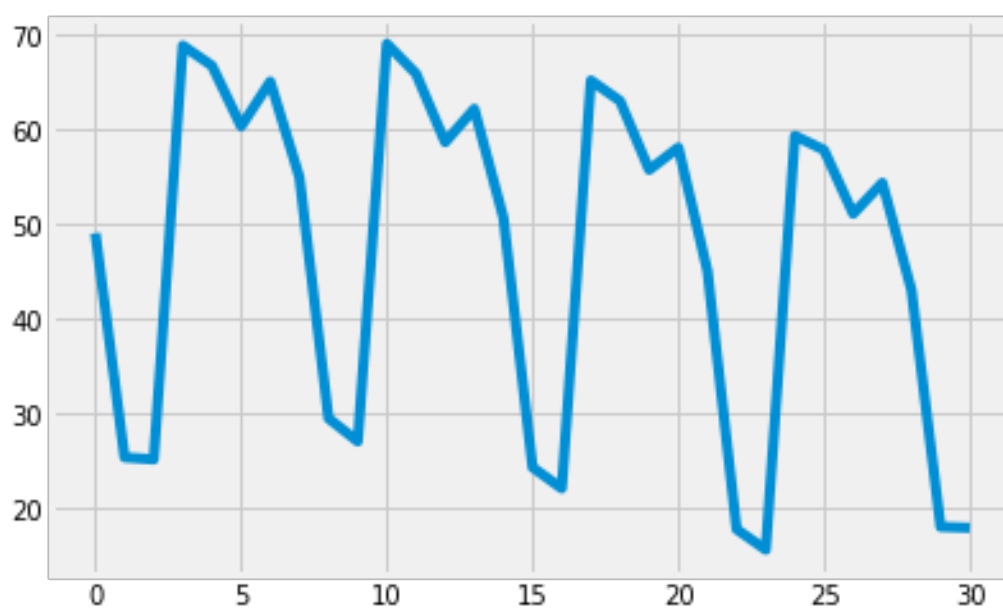
Predicción : Train set + test set con uncertainty // + con changepoints



Histórico sin predecir:



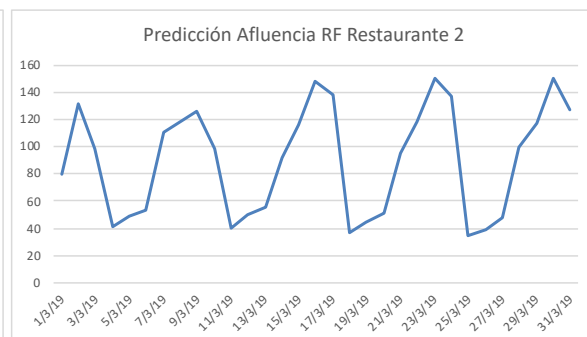
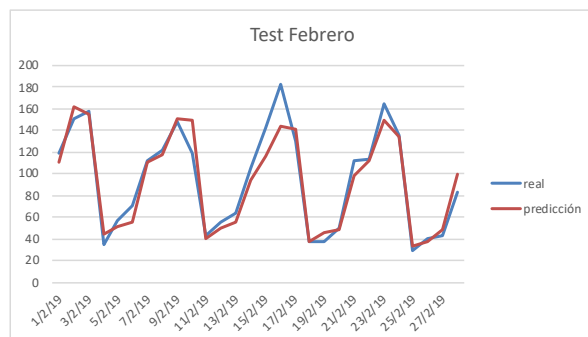
Predicción Marzo:



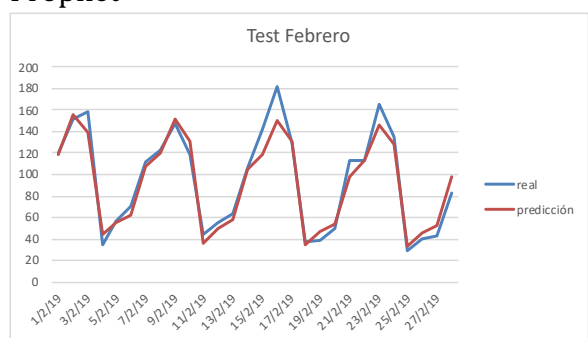
ANEXO E – Gráficas comparativas del conjunto de testeo de ambos modelos

RESTAURANTE 2:

Random Forest

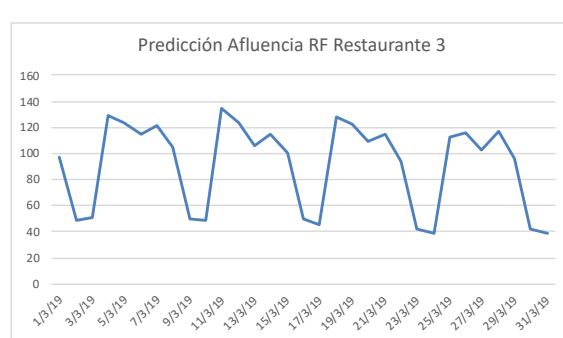
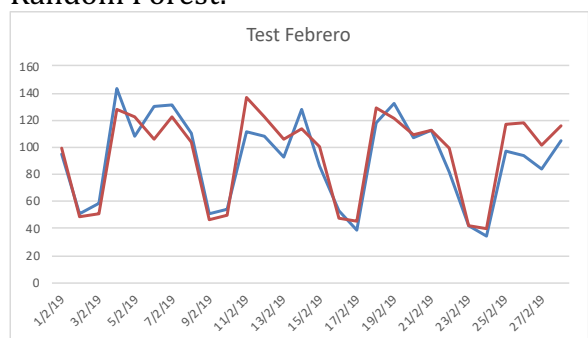


Prophet

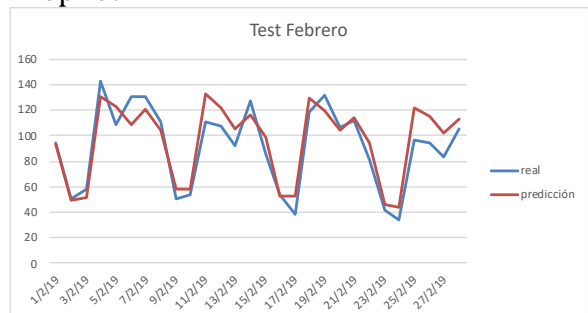


RESTAURANTE 3:

Random Forest:

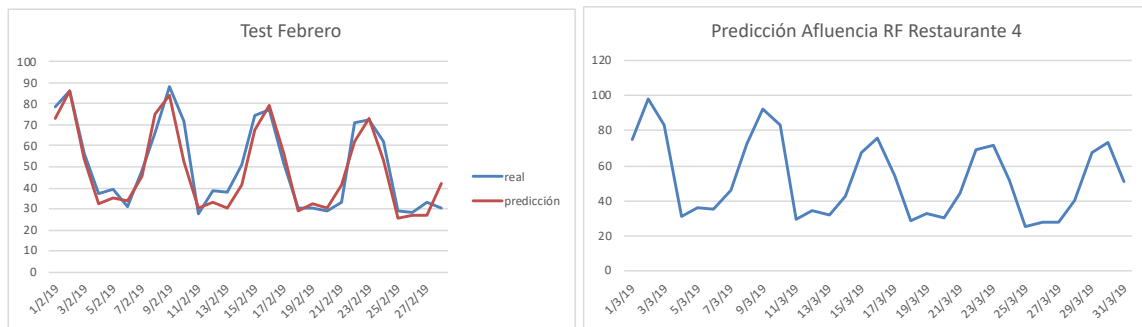


Prophet:

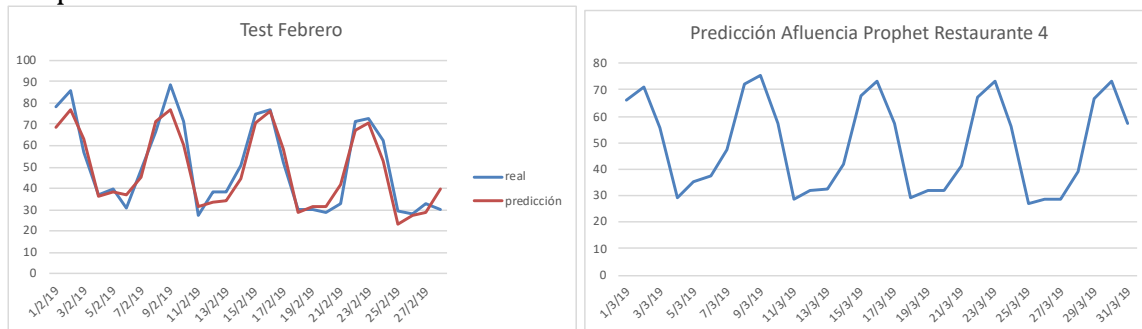


RESTAURANTE 4:

Random Forest:

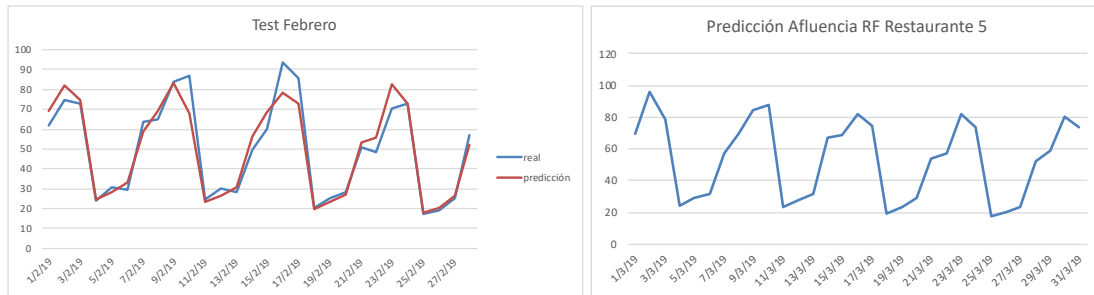


Prophet

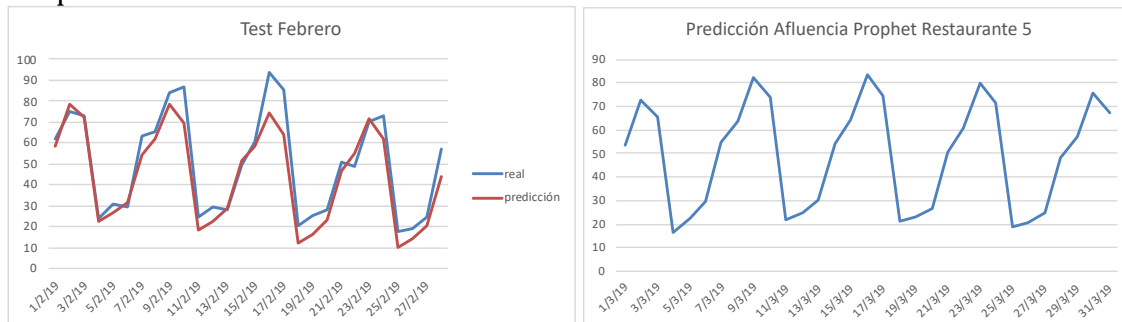


RESTAURANTE 5:

Random Forest:

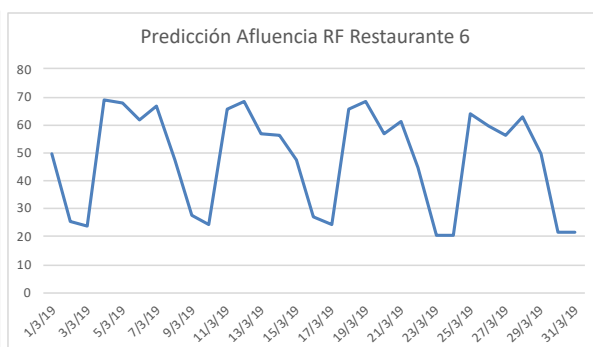
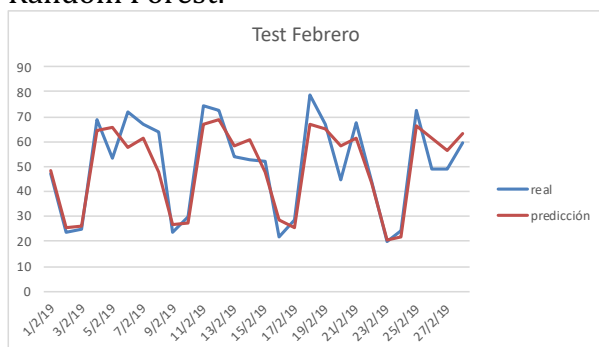


Prophet:

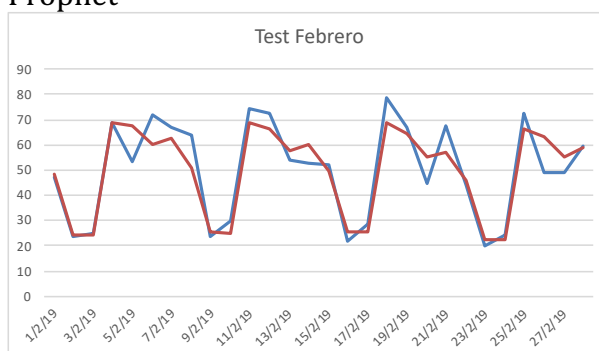


RESTAURANTE 6:

Random Forest:

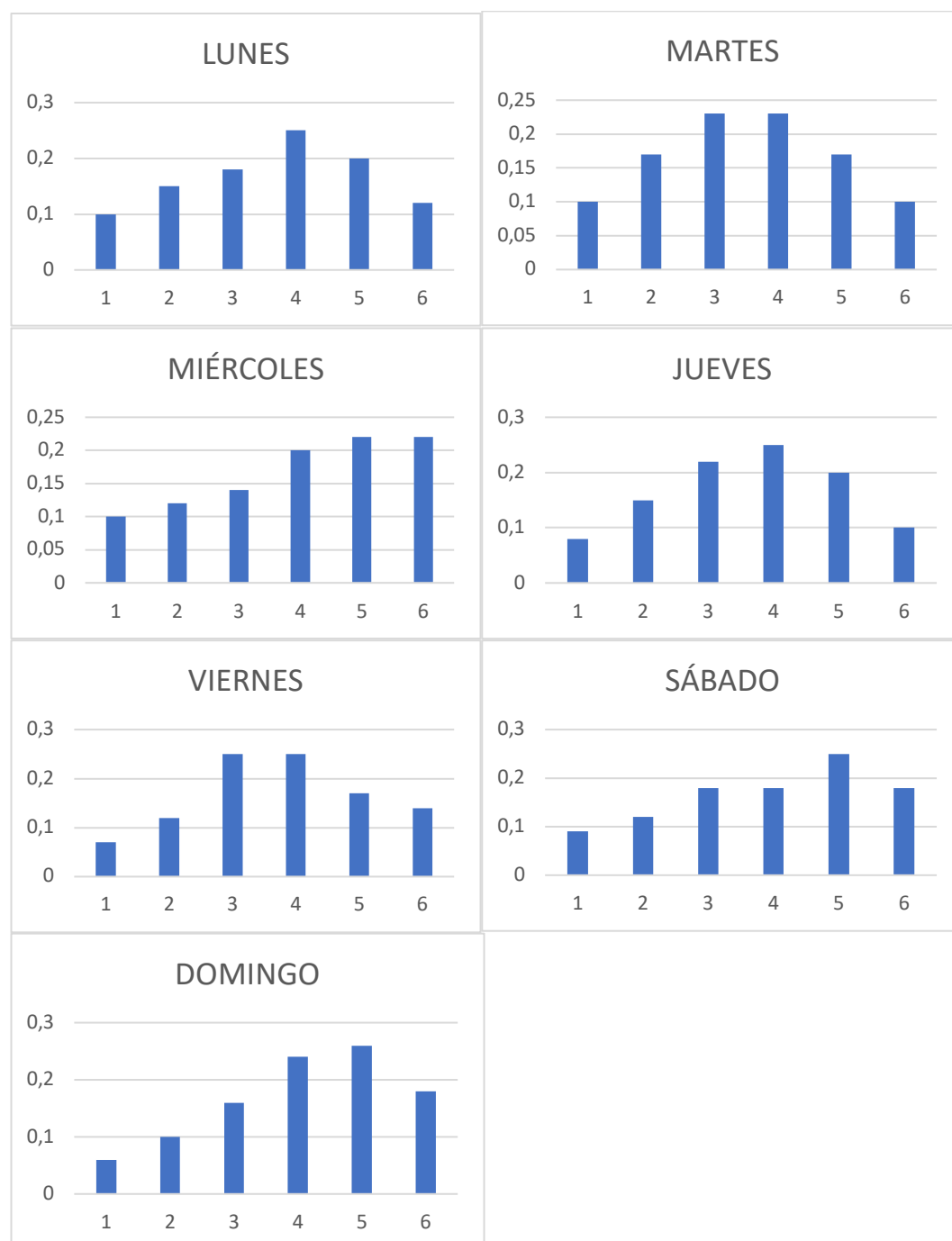


Prophet

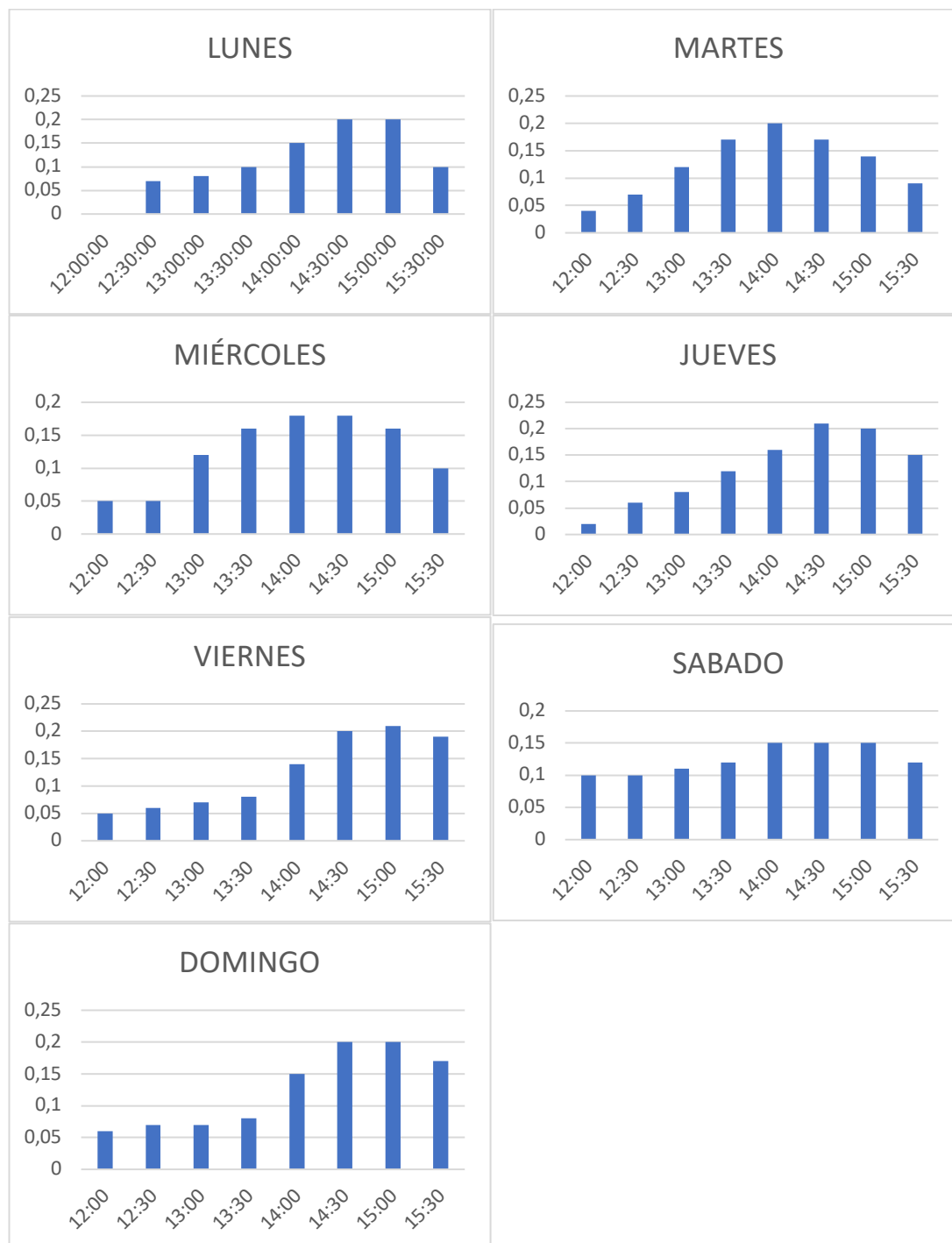


ANEXO F – Gráficas de distribución de los factores de afluencia

RESTAURANTE 2:



RESTAURANTE 3 y 6:



RESTAURANTE 5:

