OPTIMIZACIÓN BAYESIANA DE MODELOS DE NEGOCIO PARA *START-UPS*

TRABAJO DE FIN DE GRADO

BUSINESS ANALYTICS

E3 Analytics 2018-2023

Autor: Micaela Jiménez Awuapara

Director: Eduardo César Garrido Merchán

ÍNDICE DE CONTENIDOS

ABREVIATURAS

AF	Función de Adquisición	***A**cquisition **F**unction*
EI	Mejora Esperada	***E**xpected **I**mprovement*
GP	Proceso Gaussiano	***G**aussian **P**rocess*
SM	Modelo Surrogado	***S**urrogate **M**odel*

CAPÍTULO I: INTRODUCCIÓN

Optimizar una función consiste, describiéndolo de una manera muy simplista, en encontrar sus puntos máximos y mínimos, aquellos puntos en los que la función alcanza su máximo rendimiento. Existen diversos métodos matemáticos y computacionales para optimizar funciones; la aplicación de uno u otro varía dependiendo de las características de las funciones a optimizar, los objetivos, los recursos disponibles y demás variables que rodean un problema matemático. En este trabajo se estudiará la optimización bayesiana. Esta clase de métodos de optimización sirve para optimizar una función de *caja negra*.

¿Qué implica una función de caja negra y cómo se optimiza con el método bayesiano? Una función de caja negra cumple tres condiciones. En primer lugar, se da este tipo de función cuando la expresión analítica es desconocida, es decir, que las gradientes de la misma son desconocidos a su vez. No se puede representar explícitamente la función. En segundo lugar, se da cuando el resultado de la función no es determinista. Esta situación se da cuando los factores presentes en el entorno de la función producen una evaluación potencialmente ruidosa. En tercer y último lugar, cuando la función es muy costosa de optimizar. La metodología bayesiana crea una solución para esta dificultad de optimizar funciones de caja negra sin gradientes y con ruido.[1]

La optimización bayesiana es una estrategia para transformar un problema que no se puede resolver a una serie de problemas que sí se pueden resolver. La optimización en sí es una metodología que realiza la optimización global de una función multimodal de caja negra. La función se desconoce, y la optimización se centra únicamente en la optimización del rango de los hiperparámetros.

En este trabajo de investigación, se realizará una simulación de una optimización bayesiana aplicada a la optimización de un modelo de negocio de una *start-up*, planteando el problema desde una perspectiva de priorización de beneficios y la maximización de ellos. Para realizar dicha simulación o experimento, se debe construir una función que simule los beneficios de una empresa de reciente creación, ya que se trata de un experimento ilustrativo, de no ser el

[1] González, J., "Introduction to Bayesian Optimization", *Gaussian process summer school*, Sheffield University, 2017 (disponible en http://gpss.cc/gpss17/slides/gpss_bayesopt2017.pdf)

caso, se desconocería la expresión analítica y se tendría únicamente un *input* y un *output*. Para elaborar la función a optimizar, se determinará en primer lugar unas variables a tomar en consideración, y se fijarán, en segundo lugar, los rangos de estas variables. Una vez obtenido lo anterior, se procederá a ejecutar un código de optimización bayesiana sobre la función.

Este trabajo se dividirá en una exposición del estado del arte de la optimización bayesiana, una introducción a los objetivos, hipótesis, restricciones y asunciones del experimento de simulación, un *toy experiment,* una exposición del experimento ilustrativo y una conclusión sobre sus resultados.

CAPÍTULO II: ESTADO DEL ARTE

La búsqueda de hiperparámetros a través de la metodología de la optimización bayesiana está adquiriendo cada vez mayor popularidad. Cantidad de sectores están empleando estas técnicas. La ventaja que tiene este tipo de optimización es que se puede aplicar a cualquier función de caja negra, y estas pueden darse en prácticamente cualquier sector.

Se explicará brevemente a continuación cómo funciona esta metodología y algunos de los usos que se le ha dado en nuestra sociedad.

1. METODOLOGÍA DE OPTIMIZACIÓN BAYESIANA

Para poder optimizar una función de caja negra, la optimización bayesiana se estructura en una serie de procesos.

La función objetivo del problema se sustituye por un modelo surrogado probabilístico de la misma. Normalmente, se emplea un proceso gaussiano (GP) para llegar a dicho modelo surrogado. Con el GP, se computa una distribución predictiva del objetivo en aquellas regiones en el espacio donde no se ha evaluado aún la función. A continuación, basándose en la información obtenida, la optimización bayesiana computa, en cada iteración, una función de adquisición (AF). Esta estima, para cada punto de *input* en el espacio, la utilidad esperada evaluando la función objetivo en dicho punto. El punto que maximiza la AF se emplea para la evaluación iterativa. Este punto maximiza el balance entre explorar áreas desconocidas y explotar soluciones prometedoras. En otras palabras, la AF guía la búsqueda del punto óptimo. El GP se perfila con el resultado de la evaluación del punto empleado y, una vez finalizado dicho proceso, se repite el método bayesiano iterativamente hasta agotar el presupuesto. Finalizado el proceso, el método proporciona la recomendación final.[2]

---

[2] Garrido Merchán, E. C., "Advanced Methods for Bayesian Optimization in Complex Scenarios", Escuela Politécnica Superior Computer Science Department, Universidad Autónoma de Madrid, 2021 (disponible en https://repositorio.uam.es/handle/10486/699441)

2. APLICACIONES REALES DE OPTIMIZACIÓN BAYESIANA

El go es un juego de tablero que se remonta aproximadamente a unos 3.000 años, originario de China o del Himalaya. Con cierto parecido al ajedrez, el go es un juego de estrategia. Se dice que la complejidad de un juego de go se equipara a la de jugar cuatro partidas de ajedrez de manera simultánea. Las reglas del go son muy sencillas, y llevan sin cambiar desde que el juego se introdujo a la Corte Imperial de Japón en el siglo VIII. En este juego, la capacidad analítica es crucial, pero también lo es la intuición.[3] De ahí la complejidad de programar una máquina para que gane en este juego.

DeepMind es una empresa que nace en el 2010 con el objetivo de aplicar la inteligencia artificial al campo de los juegos tradicionales. Con un enfoque interdisciplinario, combinaron el *machine learning*, la neurociencia, la ingeniería, las matemáticas, la simulación y la infraestructura de computación para crear el primer programa capaz de ganar a un jugador profesional de Go: AlphaGo. En el 2015, la máquina ganó al campeón de Europa en una partida de 5-0. En enero de 2017, se lanzó una versión mejorada de AlphaGo llamada Master. La plataforma *online* ganó sesenta juegos seguidos a jugadores profesionales.[4] Todos estos logros fueron posibles al combinar las estrategias con el uso de la optimización bayesiana.

Durante el proceso de desarrollo de AlphaGo, los múltiples hiperparámetros del juego fueron optimizados con optimización bayesiana múltiples veces (como bien se explicó en la sección anterior, el proceso se repite tantas veces como permita el presupuesto). El proceso consiguió una mejora sustancial en el juego. AlphaGo se desarrolló en dos etapas: una primera etapa de entrenamiento con redes neuronales y una segunda etapa de juego con búsqueda de árboles Monte Carlo (MCTS). Ambas etapas contienen múltiples hiperparámetros. La optimización bayesiana se centró en la segunda etapa. Como se mencionó anteriormente, el go es un juego que llama a la estrategia, pero también a la intuición. En la etapa de entrenamiento, había la suficiente estrategia como para ajustar y modificar las redes neuronales, pero en la etapa del juego, no había suficiente conocimiento de cómo ajustar AlphaGo.

---

[3] Checa, A., "Historia del Go", *Club Go Madrid,* 2007 (disponible en https://www.clubgomadrid.org/go.php)

[4] DeepMind, "About Us", *AlphaGo* (disponible en https://www.deepmind.com/about)

Con la optimización bayesiana se ajustaron múltiples hiperparámetros en el juego, Componentes como aquellos que regían la función de exploración UCT, aquellos límites de expansión de nodos, aquellos involucrados en la distribución de la implementación de MCTS… La optimización bayesiana no sólo redujo el tiempo y el esfuerzo que requeriría un ajuste manual de los hiperparámetros, sino también incrementó el éxito de la máquina en el juego real. En la imagen mostrada a continuación, se muestra el proceso de las seis primeras iteraciones de optimización bayesiana.

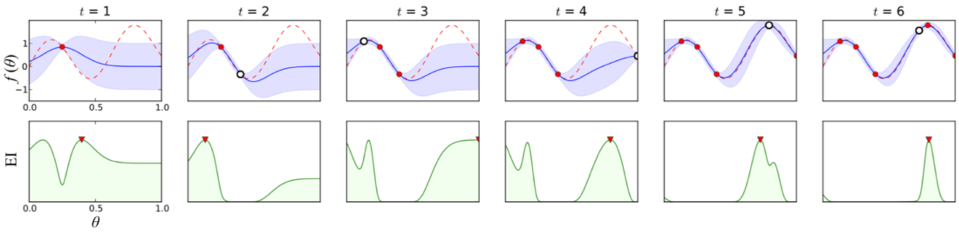

Imagen 1: ilustración unidimensional de la optimización bayesiana con proceso gaussiano y la mejora esperada en la función de adquisición en las seis primeras iteraciones. La media del proceso gaussiano en azul y la función real desconocida en rojo.

En el sector biosanitario, de acuerdo con la prestigiosa revista *Nature*, el diseño óptimo de experimentos siempre se realiza en el marco de una optimización bayesiana que lleva a la exploración más eficiente del espacio de diseño. El impacto de esta metodología en algunos campos de la ciencia es sustancial. En particular, en el campo del descubrimiento autónomo de materiales. La optimización bayesiana consigue buscar máximos de propiedades deseadas en materiales.

Por otro lado, el sector del arte también está introduciendo la optimización bayesiana. En los últimos años, se está desarrollando el arte robótico, es decir, el arte creado por inteligencia artificial. Aunque aún se trata de un campo en proceso de desarrollo, cada vez se está conociendo más del mismo. Un trabajo de investigación de la Universidad Politécnica de Valencia ha analizado cómo la optimización bayesiana se puede emplear para optimizar, por ejemplo, el número de píxeles que debe tener el arte creado por estas máquinas[5].

---

[5] Zakharyan, Elizabet, “Estudio de técnicas de optimización bayesiana aplicadas a operaciones de pintura con manipuladores robóticos”, Universidad Politécnica de Valencia, 2021 (disponible en: https://riunet.upv.es/bitstream/handle/10251/174640/Zakharyan%20-

Por último, cabe mencionar que cada vez se encuentran más usos para esta metodología en el sector financiero, en particular dentro del mundo de la optimización de portafolios en gestión de patrimonios. La optimización bayesiana se aplica en el contexto de gestión de patrimonio cuantitativo. Se tiene un problema de asignación de activos que es una estrategia de seguir patrones, y luego se resuelve con optimización bayesiana, buscando los hiperparámetros óptimos para esa estrategia de seguir patrones.[6]

Estos son sólo algunos ejemplos de las utilizaciones que se le están dando a esta metodología en el mundo de hoy, sin embargo, hay muchas otras no descritas aquí y aún más que están aún por descubrirse.

%20Estudio%20de%20tecnicas%20de%20optimizacion%20bayesiana%20aplicadas%20a%20operaciones%20de%20pintura%20con%20....pdf?sequence=1)

[6] Gonzalvez, J., Lezmi, E., Roncalli, T., Xu, J., "Financial Applications of Gaussian Processes and Bayesian Optimization", Amundi Asset Management, Paris, 2019 (disponible en: http://thierry-roncalli.com/download/Bayesian_Optimization.pdf)

CAPÍTULO III: MARCO TEÓRICO

1. OBJETIVOS

El objetivo de esta tesis de investigación es simular la optimización, a través de la metodología de la optimización bayesiana, de un modelo de negocio de una *start-up*. A través de esta simulación se intentará encontrar la manera más eficaz de emplear los recursos de dicha empresa para obtener los máximos beneficios.

Las *start-ups* suelen levantar rondas de financiación, empezando por ángeles y seguido de fondos de capital venture y, a través de dichas rondas, desarrollan su negocio. La clave del éxito de estas empresas, desde un punto de vista operativo-financiero, es el empleo eficaz y eficiente de dicho financiamiento. Por ello, resulta de vital importancia encontrar formas de apoyarse en los datos y en las matemáticas detrás de la optimización de modelos para generar el mayor rendimiento y potenciar a las empresas de reciente creación.

A continuación, los objetivos de la tesis de investigación:

- Objetivo #1: Realizar una memoria del trabajo realizado en el TFG.
- Objetivo #2: Realizar un experimento ilustrativo de optimización de un negocio de *start-up* con optimización bayesiana.
- Objetivo #3: Obtener resultados de dicho experimento ilustrativo que se puedan extrapolar a *start-ups* con variables en el modelo de negocio distintas.

2. RESTRICCIONES

Las restricciones de cara a este experimento ilustrativo se basan principalmente en el tiempo y en los recursos.

El tiempo disponible para este TFG es limitado. Únicamente se ha podido empezar con la investigación de este trabajo en el mes de enero y debe ser entregado a finales de abril, por lo que no se dispone de demasiado tiempo.

En segundo lugar, los recursos. Se dispone de un ordenador personal, un Macbook Pro de 2017 con 8GB de memoria interna y un procesador de 3,1 GHz Dual-Core Intel Core i5. No es un ordenador profesional en el que se puedan realizar programaciones complejas e iteraciones de modelos que requieran un mayor procesador. Por ello, cómo se explicará más adelante, se ha recurrido a un procesador externo: Google Colab.

En último lugar, cabe mencionar que la limitación principal es el conocimiento previo sobre la materia. A pesar de la formación para este trabajo, todo el conocimiento para realizarlo viene de dicha formación. No se dispone de conocimientos previos sobre la optimización bayesiana y esto limita la capacidad de tunear el modelo o la optimización para hacerla más precisa o compleja.

3. HIPÓTESIS

Hipótesis 1: La función de beneficios de una *start-up* es modelable con un proceso gaussiano.

Hipótesis 2: Los resultados obtenidos con la metodología de optimización bayesiana serían superiores a aquellos resultados obtenidos con la metodología de búsqueda aleatoria en cuanto a rendimiento obtenido de la función.

4. ASUNCIONES

Asunción 1: Cada experimento ilustrativo contempla sólo cinco variables que sirven para hacer una buena modelización de un problema empresarial.

Asunción 2: Distintos sectores de negocios de *start-ups* exhiben el mismo comportamiento.

Asunción 3: Para efectos del experimento ilustrativo, las unidades de ventas de la *start-up* coincidirán con las unidades de producción.

CAPÍTULO IV: DISEÑO TEÓRICO DE LA METODOLOGÍA

En esta sección se detallará el proceso que se sigue de optimización bayesiana, reflejado en el código del experimento. Detallaré en orden la lógica detrás de la programación del mismo, principalmente del orden y de la función de los comandos.

En un simple diagrama, el algoritmo de optimización bayesiana consistiría en el siguiente bucle: observar los puntos de datos, construir el modelo surrogado basado en un proceso gaussiano (GP), maximizar la función de adquisición, muestrear el siguiente punto y añadir punto a conjunto de valores observados.

1. LA FUNCIÓN OBJETIVO

Toda optimización bayesiana parte de una función objetivo desconocida, ergo el que se le denomine optimización de una función de caja negra. Sin embargo, para efectos de este trabajo de investigación, y ya que se trata de un experimento simulado, se definirá en primer lugar una función objetivo. Esta será desconocida de cara a la optimización, pero definida para el experimento.

2. CONJUNTO DE FUNCIONES

Una vez establecida la misma, se generan puntos aleatorios en el espacio. A este conjunto se le denomina conjunto de entrenamiento. Esto servirá para entrenar al modelo de optimización. Luego, se computa la función latente de la función objetivo. Esto se realiza con el comando *torch.unsqueeze*. Este comando devuelve un tensor nuevo con una dimensión de tamaño uno en la posición especificada. El tensor nuevo comparte un mismo tensor, o *input*, subyacente.[7] Una vez computada dicha función, se calcula el valor observado que maximiza la misma. Este proceso se condensa en una única función que genera los tres retornos mencionados anteriormente: los puntos aleatorios, la función latente y el valor que maximiza la función latente. Con esta función anterior de condensación, se empieza el proceso iterativo de optimización bayesiana.

---

[7] Meta, "Torch.unsqueeze", *PyTorch*, 2022 (disponible en https://pytorch.org/docs/stable/generated/torch.unsqueeze.html)

3. LÍMITES Y MODELO

Antes de comenzar el proceso iterativo, se deben ajustar los límites. Son los límites del tensor, *torch.tensor*. Dentro de ellos, se buscará optimizar la función de adquisición. Se trata de los límites dentro de los cuáles se ajustarán los hiperparámetros. Una vez establecidos, se determina el modelo que se utilizará. En la mayoría de las situaciones en las que se emplea la optimización bayesiana, se utiliza un GP. Es lo que se utilizará en este trabajo. De igual manera, en cuanto a la computación de los hiperparámetros, se hará la misma utilizando la verosimilitud logarítmica marginal exacta.

Una vez declarado el modelo y el proceso a seguir detallado anteriormente, se ajustan los puntos aleatorios anteriores con el GP y se determinan los hiperparámetros con el proceso de verosimilitud logarítmica marginal exacta. El resultado, *output*, de lo anterior muestra la función de covarianza empleada por el proceso gaussiano por defecto y sus hiper-hiperparámetros.

a) Proceso Gaussiano

La optimización bayesiana se utiliza para conseguir el extremo x^* de una función de caja negra $f(x)$: $x^* = \arg min_{x \in \chi} f(x)$, asumiendo la minimización. Para poder realizar este proceso de hallar el extremo, se depende de un modelo probabilístico surrogado de la función de caja negra, la función *latente*. Esta normalmente se basa en un proceso de probabilidad gaussiana. Se emplea la distribución predictiva gaussiana $p(f(x)|\mathcal{D})$ para generar una superficie de respuesta que mida la utilidad o mejora esperada en cada punto $x \epsilon \chi$. Esta superficie de respuesta es la AF (véase explicación de la sección D. Función de adquisición).[8]

El GP asume en primer lugar $y_i = f(x_i) + \varepsilon_i$, donde ε_i es el sonido gaussiano aditivo $\varepsilon_i \sim \mathcal{N}(0, \sum \boldsymbol{I})$, de tal manera que la $\boldsymbol{I}$ es la matriz de identidad, y_i es el valor observado de la

---

[8] Garrido Merchán, E. C., "Advanced Methods for Bayesian Optimization in Complex Scenarios", Capítulo 3, Escuela Politécnica Superior Computer Science Department, Universidad Autónoma de Madrid, 2021 (disponible en https://repositorio.uam.es/handle/10486/699441)

función objetivo $f(x_i)$ evaluada en el punto x_i. La función objetivo está corrompida por el sonido gaussiano aditivo descrito, por lo que $f(x_i)$ es el valor de la función latente que no observamos. Para cada punto de **x** en el espacio de input de χ perteneciente al rango de la función objetivo, el GP calcula la distribución gaussiana para el valor de $f(x)$ asociado a la **x**. Los valores potenciales de $f(x)$ se especifican con la distribución predictiva del GP, $p(f(x)|\mathcal{D})$, donde $\mathcal{D} = \{(x_i, y_i)|i = 1, \dots, N\}$ es un conjunto de datos de observaciones anteriores.[9]

La distribución predictiva gaussiana queda determinada por la función de covarianza gaussiana, la cual introduce las asunciones que se hacen sobre la función objetivo. Las funciones de covarianza contienen hiperparámetros que condicionan la distribución predictiva. De ahí la importancia de estimar correctamente los hiperparámetros para que haya un ajuste preciso de la función objetivo.

b) Verosimilitud logarítmica marginal

La verosimilitud logarítmica marginal se describe con la siguiente expresión analítica:

$$\log p(y|x, M_i) = -\frac{1}{2}y^{\intercal}K^{-1}y - \frac{1}{2}\log|K| - \frac{n}{2}\log(2\pi)$$

Se trata de una combinación entre un término de ajuste de los datos, la parte de la expresión analítica anterior marcada en azul, y una penalidad de complejidad, la parte de la expresión marcada en verde. El determinante $|K|$ representa cuanto ocupan los parámetros del modelo. Si el espacio que ocupa es demasiado grande, se puede estar haciendo un *overfitting*, o sobreajuste, sobre el modelo. Este es un efecto que no se desea. Por ello, se le determina a ese término como penalidad de complejidad. Por otro lado, el uso del logaritmo implica el productorio de todas las variables aleatorias. Se trata de un productorio de probabilidades. Para encontrar los valores de los parámetros que mejor se ajusten al modelo, se debe encontrar aquellos valores con mayor probabilidad, ya que serán los que mejor se ajusten. Todas esas probabilidades se juntan en dicho productorio que, al ser de tal gran dimensión, se muestran con una expresión logarítmica que hace que su dimensión sea manejable.

---

[9] Garrido Merchán, E.C., Capítulo 2.

El aprendizaje en un GP implica encontrar dos puntos principales: la forma de la función de covarianza y los (hiper)parámetros desconocidos θ. Esto se puede conseguir al optimizar la función de verosimilitud marginal:[10]

$$\frac{\partial \log p(y|x,\theta,M_i)}{\partial \theta_j} = -\frac{1}{2} y^\intercal K^{-1} \frac{\partial K}{\partial \theta_j} K^{-1} y - \frac{1}{2} trace(\mathrm{K}^{-1} \frac{\partial K}{\partial \theta_j})$$

Este proceso de computar la función anterior lo realiza el comando de BoTorch *ExactMarginalLogLikelihood*, que se importa de la librería *gpytorch.mlls.exact_marginal_log_likelihood.*

4. FUNCIÓN DE ADQUISICIÓN

El siguiente paso es establecer la función de adquisición (FA) que se va a computar utilizando la distribución predictiva del anterior GP en todo el espacio de *input*. Para ello se utilizará la mejora esperada, la función dentro de BoTorch de *qExpectedImprovement*.

La función de adquisición (AF) es un elemento característico de la optimización bayesiana. Al desconocer la función objetivo, no se puede maximizar. En vez de maximizar la función objetivo $f(x)$, se maximiza la FA, la cual es mucho más fácil y menos cara de optimizar. El modelo surrogado (SM) se obtiene con el GP explicado anteriormente, un proceso que logra ajustar los puntos observados y cuantificar la incertidumbre de las áreas no observadas. El SM es la mejor manera de aproximar la desconocida función de caja negra $f(x)$. La AF observa el SM y determina qué áreas del dominio de $f(x)$ tienen una probabilidad lo suficientemente contundente como para que merezca la pena explorarlas – así empleando eficientemente los recursos disponibles – de tal manera que las áreas que merece la pena explorar son aquellas en las que la función desconocida es óptima o áreas que aún no se han mirado. La AF asume un

[10] Las expresiones analíticas y la explicación provienen de: Rasmussen, C. E., “Gaussian Processes for Machine Learning”, *Carlos III*, 2006 (disponible en: https://www.tsc.uc3m.es/~fernando/l1.pdf)

valor alto en dichas áreas. Por el contrario, en aquellas áreas del dominio en las que la función objetivo es subóptima o áreas que no se hayan muestreado, tendrán un valor bajo.[11]

a) Probabilidad de mejora

Si suponemos que la función a maximizar es $f(x)$, y la mejor solución que se tiene hasta este momento es x^*, la mejora se puede definir como:

$$I(x) = \max\left(f(x) - f(x^*), 0\right)$$

De esta manera, si el nuevo valor de x tiene un valor asociado $f(x)$ que es menor que $f(x^*)$, entonces $f(x) - f(x^*)$ será negativo. No se está mejorando, no se está encontrando un valor superior de la función. Sin embargo, si el nuevo valor de $f(x)$ es superior que la estimación que hay ahora, la diferencia será positiva. En este caso, $I(x)$ devuelve la diferencia, en otras palabras, cuánto mejorará la función al evaluarla en el punto nuevo.

En una AF de probabilidad de mejora, para cada punto candidato de x, se asigna una probabilidad de $I(x) > 0$, es decir, que $f(x)$ sea superior al valor de la función con el actual mejor candidato $f(x^*)$ – ello implica que la función, en caso encontrar un candidato de x que incremente el valor de la función, estaría más optimizada con este nuevo valor que con el anterior óptimo. Esta explicación teórica, se relaciona con el GP ya que en cada punto candidato existe una distribución gaussiana asociada al mismo. Por tanto, en el punto x, el valor de la función $f(x)$ está muestreado desde una distribución normal con media $\mu(x)$ y una varianza $\sigma^2(x)$:

$$f(x) \sim \mathcal{N}(\mu(x), \sigma^2(x))$$

Si se emplea la estadística tradicional para estandarizar, se hace una reparametrización y se rescribe la función de mejora. Si $z \sim \mathcal{N}(0,1)$, entonces la fórmula se rescribe como:

[11] Ekamperi, S., "Acquisition functions in Bayesian Optimization", *Let´s talk about science*, 2021 (disponible en https://ekamperi.github.io/machine%20learning/2021/06/11/acquisition-functions.html)

$$I(x) = \max(f(x) - f(x^*), 0) = \max(\mu(x) + \sigma(x)z - f(x^*), 0)\ z \sim \mathcal{N}(0,1)$$

Sintetizando, x es un valor del que se desea ver si merece la pena evaluar en $f(x)$. Por tanto, se le asigna un valor de $I(x)$ a él. Sin embargo, el valor de $I(x)$ se muestrea con una distribución normal $\mathcal{N}(\mu(x), \sigma^2(x))$, de tal manera que se calcula como:

$$PI(x) = \Pr(I(x) > 0) \leftrightarrow \Pr(f(x) > f(x^*))$$

b) Mejora esperada

La probabilidad de mejora considera únicamente la probabilidad de mejorar la estimación actual, el valor mejor observado hasta el momento, pero no introduce como factor la magnitud de la mejora. En esta característica se distingue la A de mejora esperada, *expected improvement*. En vez de calcular la mejora, $I(x)$, que es una variable aleatoria, se calcula la mejora esperada (EI). Esto es el valor esperado de $I(x)$:

$$EI(x) \equiv \mathbb{E}[I(x)] = \int_{-\infty}^{\infty} \max(f(x) - f(x^*), 0)\, \varphi(z) dz$$

Donde $\varphi(z)$ es la función de densidad de probabilidad de la distribución normal $\mathcal{N}(0,1)$, es decir, $\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$.[12]

Sin entrar a mayor detalle del cálculo matemático detrás de la resolución de esta ecuación, el comando *qExpectedImprovement* dentro de la librería *BoTorch.Acquisition.Monte_Carlo*.

5. OPTIMIZACIÓN DE LA FUNCIÓN DE ADQUISICIÓN

Con la función de adquisición calculada, se procede a optimizar la misma. Se intenta encontrar el óptimo global. Para ello, se condensa todo el código de la iteración en una única función que se compone en forma de bucle de optimización bayesiana. Esta retorna los puntos candidatos

[12] Ekamperi, S., "Acquisition functions in Bayesian Optimization", *Let´s talk about science*, 2021 (disponible en https://ekamperi.github.io/machine%20learning/2021/06/11/acquisition-functions.html)

a óptimo global. En esta función se debe indicar, a la hora de ejecutarla, el número de iteraciones que la función debe hacer. Remontando a la explicación inicial de la optimización bayesiana, esta se aplica hasta que se cumpla el presupuesto. Una vez ejecutada dicha función con el número de bucles deseado, se le solicita que se obtengan los puntos que observan mejor resultado en la optimización. De esta manera, se puede ver si la optimización ha tenido éxito o no.

En cada iteración, se encuentra la x que maximiza la AF. Con esta información, se identifica la próxima gran conjetura para probar en f. Aquella que se probará en la próxima iteración.

6. GRÁFICOS OPCIONALES

De manera adicional, se puede considerar dibujar la función de adquisición con su punto máximo, el punto que se sugiere para ser evaluado en la siguiente iteración. Se puede hacer este dibujo en cada iteración. Si se trata de una única variable, el dibujo es simple, si se tienen más variables, se puede elegir una para mostrar (es el caso del experimento simulado principal de este trabajo). También se puede mostrar en un gráfico la media predictiva del proceso gaussiano, su desviación típica y su distribución predictiva para todo el espacio de input. Estos adicionales se pueden introducir en el bucle iterativo de optimización bayesiana principal (es el caso del *toy example*.

CAPÍTULO V: DETALLES DE LA IMPLENTACIÓN

Para la implementación de la optimización bayesiana en el experimento de simulación, se ha empleado como plataforma de programación un cuaderno de Jupyter. Esta plataforma es un *open-source software* que proporciona una plataforma interactiva de computación. En este caso, se ha utilizado un cuaderno de Jupyter en Google Colab.

Google Colab permite hacer un uso de una plataforma de programación sin necesidad de instalar la misma en el ordenador en el que se utiliza. Esto permite que el ordenador no se sobre-trabaje o que el ordenador no supere sus capacidades. Además, protege el trabajo ya que tiene una función de autoguardado. En este caso, se programará en Python. Este es el idioma de programación más conocido del mundo y es el que principalmente se emplea para optimización bayesiana.

La optimización bayesiana en sí se realizará a través del uso de la librería BoTorch. Esta librería, desarrollada por la compañía Meta, forma parte del ecosistema de programación PyTorch, y está específicamente diseñada para ser empleada en Python.

BoToch proporciona una sencilla y modulable interfaz de usuario para componer primitivas de optimización bayesiana, incluyendo en ello los modelos probabilísticos, las funciones de adquisición y los optimizadores. Esta herramienta, a su vez, mejora significativamente la eficiencia del desarrollador a la hora de utilizar funciones de adquisición cuasi-Monte-Carlo. Esto permite que la forma de implementar nuevas ideas sea directa, sin necesidad de imponer restricciones en las asunciones de la función objetivo.[13]

[13] Meta, "Introduction", *BoTorch* (disponible en https://botorch.org/docs/introduction.html)

CAPÍTULO VI: TOY EXAMPLE[14]

De manera ilustrativa y siguiendo el proceso detallado en la sección *Diseño teórico de la metodología*, se ha realizado un experimento denominado *toy example*. En este se pretende realizar una optimización bayesiana sobre una función simple con una única variable. A continuación, los detalles de la función, del proceso y de los resultados obtenidos.

Como se mencionó anteriormente, en una situación real de aplicación de optimización bayesiana, no se conocería la función objetivo. Es uno de los requisitos para que la función sea una función de caja negra, el que sea desconocida. Sin embargo, a efectos del experimento, la función se determinará previamente. La función objetivo de este *toy example* es:

$$f(x) = e^{-(x-2)^2} + e^{-\frac{(x-6)^2}{10}} + \frac{1}{x^2+1}$$

La función anterior contiene una única variable *x*, que en este caso será la variable inversión disponible y la *y* de la función será los ingresos. Es una versión muy simplificada del experimento ilustrativo principal y aquél que se mostrará posteriormente. La función se representa de la siguiente manera:

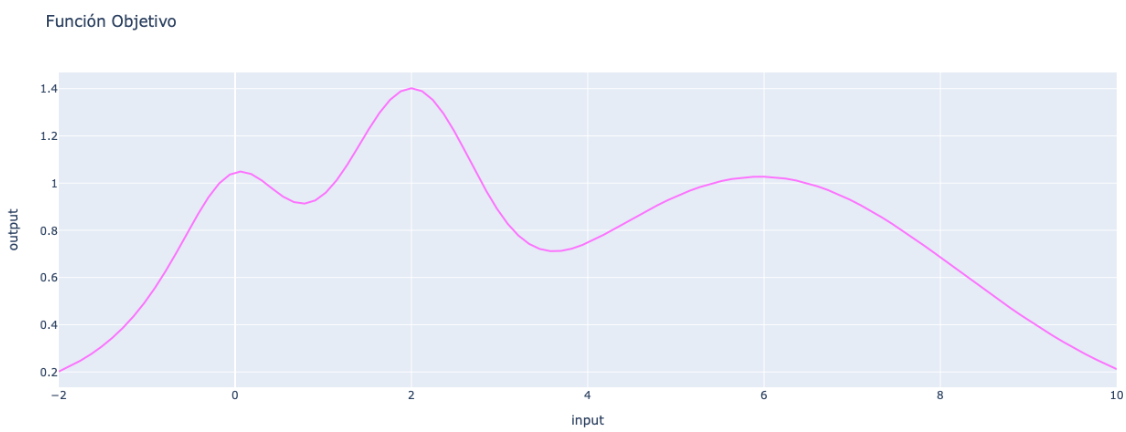

Imagen 2: representación gráfica de la función objetivo del toy example.

[14] El código del *Toy Example* se encuentra disponible en: https://colab.research.google.com/drive/1jU457fk22A6ArK5Bu7kb4fGV8ZqPXcWu?usp=sharing

Tras condensar el código para obtener la función latente, el conjunto de puntos aleatorios y el mejor valor observado, se ajustaron los límites del proceso iterativo a los siguientes:

$$f(x): x \in [0,10]$$

Con esta información, se puede ejecutar el bucle creado de optimización bayesiana en el que se busca la función de adquisición y se optimiza con el máximo de la misma.

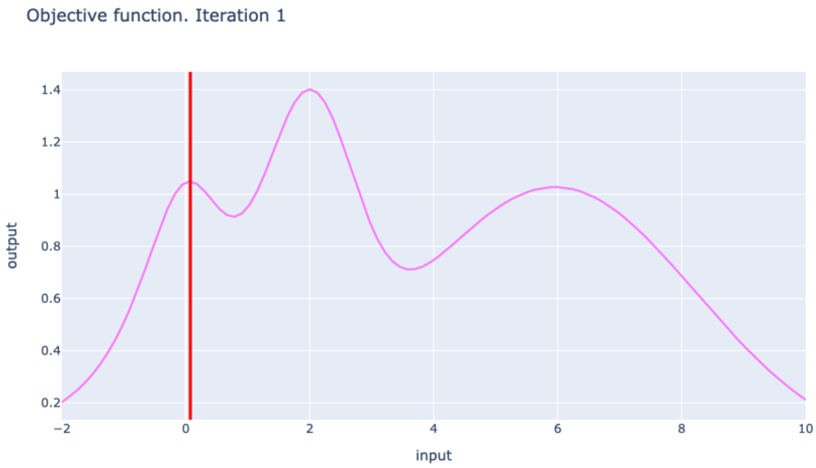

Imagen 3: función objetivo en la primera iteración de optimización bayesiana en toy example.

En la imagen anterior, se observa el mejor valor observado en la primera iteración del bucle de optimización bayesiana. Se puede observar que se trata de un valor muy cercano al 0. En este caso, se ha obtenido un máximo local. Se verá a continuación el gráfico de la mejora esperada.

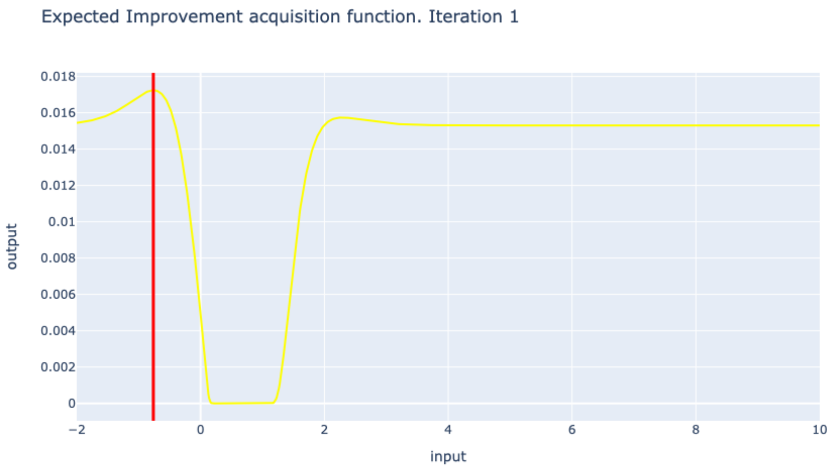

Imagen 4: mejora esperada de la función de adquisición en la primera iteración en toy example.

Se observa que la función de espera mejorada es nula en los siguientes puntos cercanos al máximo local que se ha encontrado. Esto se observa porque la iteración ya ha evaluado una sección cercana al punto máximo, por lo que la probabilidad de encontrar un máximo superior al observado en esa sección es nula. Se ignorará esta sección de cara a la siguiente iteración y se procederá a explorar secciones no exploradas de la función de adquisición. Esto se observa a su vez en la plasmación de la distribución predictiva gaussiana.

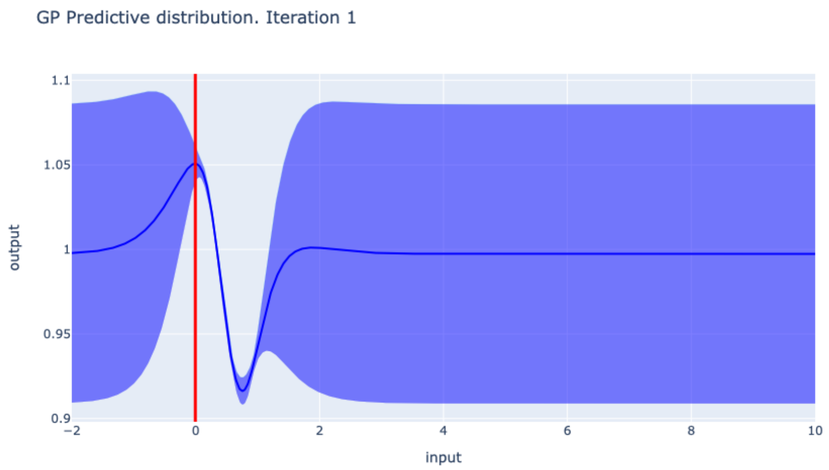

Imagen 5: distribución predictiva gaussiana en la primera iteración en toy example.

El proceso iterativo repite esta búsqueda en secciones aún no exploradas de la función donde la probabilidad de hallar un máximo es más elevada. Si se avanza a la quinta iteración de la función, el gráfico de la distribución de probabilidad gaussiana mostrará menos zonas potenciales, pues se habrán evaluado ya cinco puntos donde la función podría encontrar su máximo, y en base a los resultados, se ha explorado más o no.

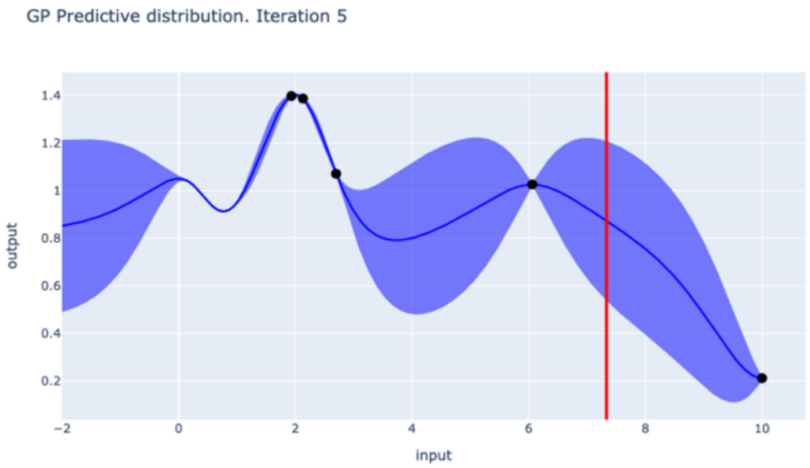

Imagen 6: distribución predictiva gaussiana en la quinta iteración en toy example.

Como observamos en la imagen anterior, quedan algunas zonas por explorar donde potencialmente se podría encontrar un punto máximo. Se continuará con iteraciones de la exploración hasta agotar el presupuesto. En este caso, el presupuesto establecido es de nueve iteraciones.

Observamos el resultado obtenido en la novena iteración.

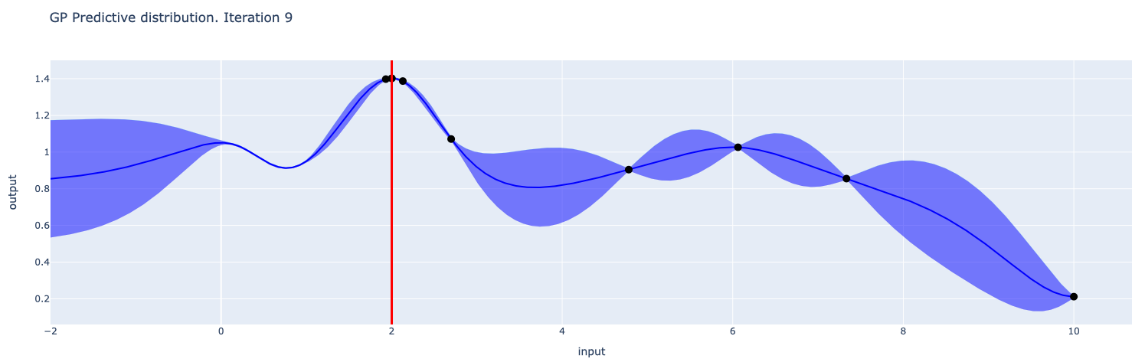

Imagen 7: distribución predictiva gaussiana en la novena iteración en toy example.

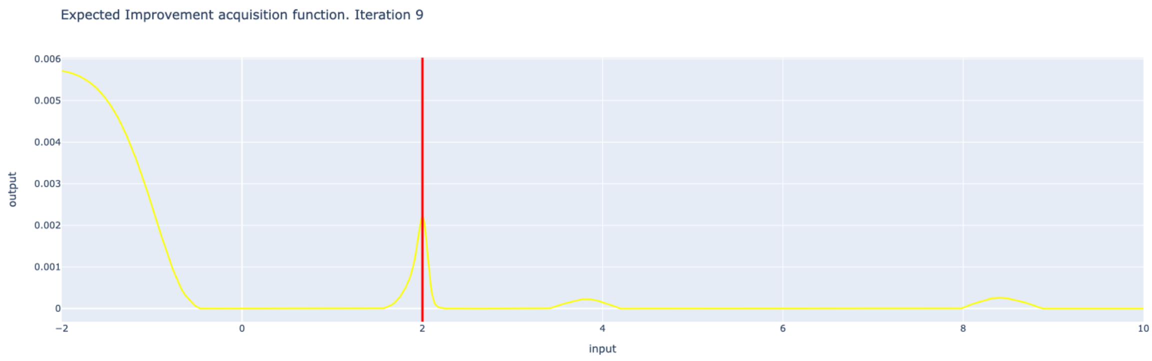

Imagen 8: mejora esperada de la función de adquisición en la novena iteración en toy example.

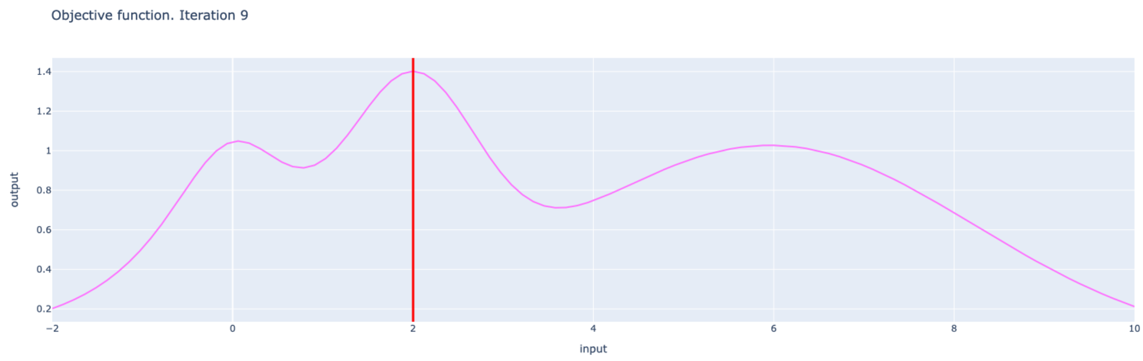

Imagen 9: función objetivo en la novena iteración de optimización bayesiana en toy example.

En este caso, la optimización bayesiana ha sido exitosa y se ha encontrado el punto óptimo de la función objetivo. Se observa que quedan zonas por explorar, ya que es imposible explorar una función en todos sus puntos. Sin embargo, se ha seguido la metodología gaussiana y se han evaluado los puntos donde mayor probabilidad había de que se encontrase el máximo. Finalmente, este se ha colocado en el valor de x de 2.

Habiendo explicado al metodología a seguir y habiéndola testado en el *toy example* anterior, se procederá a la realización del experimento ilustrativo sobre el que se basa este trabajo.

CAPÍTULO VII: EXPERIMENTO ILUSTRATIVO

Este experimento toma como referencia un modelo de negocio de una *start-up.* Estas empresas disponen de recursos limitados y deben optimizar sus beneficios de tal manera que les permita dar retorno a sus inversores y crecer. El objetivo de este experimento es aplicar la metodología de la optimización bayesiana para optimizar los beneficios de una empresa que se encuentra en sus primeros años de desarrollo. Estos beneficios se considerarán dentro de un periodo de tiempo de un año.

La función matemática a optimizar se basa en una concepción ilustrativa de la función de beneficios de una *start-up*. De manera sencilla, se representa como unos beneficios que dependen de los ingresos, los costes variables, los costes fijos, la fuerza laboral y las reservas a mantener. Mientras los costes fijos, la fuerza laboral y las reservas serán tomadas como variables independientes, los conceptos de ingresos y costes variables están relacionados entre sí y dependen de las variables unidades y precios. Los ingresos se constituyen del número de unidades vendidas multiplicado por el precio por el que se venden, mientras que los costes variables dependen únicamente de la variable unidades. Finalmente, dado lo anterior, las cinco variables a optimizar serían las unidades a producir (que cómo se mencionó en el apartado de asunciones, se asume que es el mismo que las unidades vendidas), el precio al que se deben vender dichas unidades, los empleados que debe tener la empresa, los costes fijos que se deben asumir, y, por último, las reservas que debe mantener la empresa.

$$Beneficios(precio, unidades, empleados, coste\ fijo, aprovisionemiento\_riesgos)$$

Asimismo, se debe tener en cuenta, que la restricción que aplica a esta optimización es la del capital disponible para realizar las operaciones. Esto se detallará en el apartado siguiente.

Cabe mencionar, que para las *start-ups*, la reinversión es fundamental. Es la forma que tienen estas empresas de crecer. Para efectos de este experimento ilustrativo, se considerará que los beneficios obtenidos por la empresa serán reinvertidos en su totalidad, por lo que no se consideran una variable dentro de los mismos.

Por último, enfatizar que la función objetivo empleada para este experimento es una aproximación simbólica. Se trata de una forma de enseñar la aplicación que puede tener la optimización bayesiana y como la misma es la estrategia óptima para conseguir el máximo rendimiento con una empresa de reciente creación con el capital limitado. La aproximación hecha en el presente trabajo se ha realizado a través de una inferencia basada en la teoría de modelos económicos, lo que sería un modelo hedónico prácticamente de la función objetivo.

1. CAPITAL DISPONIBLE

Como toda *start-up*, el capital disponible es la principal limitación de cara a crecer. Estas empresas, de reciente creación, dependen de fuentes externas de financiación. En particular, el capital disponible para la utilización y desarrollo de las *start-ups* viene de las denominadas rondas de financiación.

Las rondas de financiación de *start-ups* se dividen en varias etapas. Según K Fund, fondo de capital riesgo español, existen siete rondas principales de financiación. La primera es el dinero de los propios fundadores y la segunda es la famosa ronda de *family, friends and fools*, traducido a familia, amigos e insensatos, aquellos que les dan un voto de confianza a sus cercanos. En tercer lugar, se encuentra la ronda de *business angels*. Esta se trata de gente que ya se dedica a financiar ideas y proyectos de manera profesional. Ya no invierten solamente en la persona, sino en la idea y en la persona. En cuarto, quinto, sexto y séptimo lugar las llamadas series de financiación. Primero se encuentra la serie *seed*. Esta ronda es la primera en la que entran los fondos de capital riesgo o *venture capital*. En esta ronda, la financiación puede alcanzar hasta los $2 millones. Tras esta ronda ya llegan las rondas de serie A, B y C. La primera normalmente corresponde a capital riesgo genérico, la segunda a fondos de capital riesgo más específicos dirigidos a empresas más desarrolladas, y la última ronda C suele estar entre financiada por fondos de capital riesgo de etapas muy avanzadas o incluso ya fondos de *private equity*[15].

Las rondas A suelen proveer a las *start-ups* de financiación de media de $2 millones, serie B de $7.5 millones, y serie C de $20 millones.

[15] Ventura, P., "Cap tables y rondas, ¿a qué deberías aspirar?", *K Fund,* 2017 (disponible en https://blog.kfund.vc/cap-tables-y-rondas-a-qué-deber%C3%ADas-aspirar-6bdc68e6f5ea)

Para efectos del experimento ilustrativo de optimización bayesiana, consideraremos que la *start-up* en cuestión está entre la ronda *ángel* y la ronda A. Por tanto, la restricción del capital disponible por la empresa se fijará en $1 millón en el periodo de análisis de un año.

2. DINÁMICA DE LAS VARIABLES

a) Dinámica del precio (z)

i. Comportamiento de la variable

El precio del producto determina, junto con las unidades vendidas (en este caso asumiendo que son las mismas que las unidades producidas), los ingresos de la compañía.

Siguiendo la teoría económica de la dinámica de la oferta y la demanda, a mayor el precio, menor producción se venderá, y a menor el precio, mayor cantidad de ventas. Esto se da debido a que el consumidor tiene mayor poder de adquisición cuando el bien tiene un precio menor. En esta sección se podrían introducir todo tipo de teorías económicas sobre los diferentes bienes. Por ejemplo, se podría tratar de un bien inelástico. Esto implica que el bien es resistente a variaciones en el precio, normalmente ya que se trata de un bien necesario que el consumidor va a comprar de todas maneras. La variación de la cantidad demandada en una curva de oferta y demanda dependiente del precio es mínima. Algunos ejemplos de estos bienes serían los medicamentos, el agua, la luz, el pan, el tabaco, etc. Bienes necesarios para la salud, para la vivienda o susceptibles de adicción, entre otros. En este caso, la ecuación matemática representando la cantidad sería prácticamente la misma independientemente del precio. Por otro lado, existen en el lado opuesto del espectro, los bienes más elásticos del mercado. La demanda de estos depende sustancialmente del precio, teniendo una pequeña variación en el precio un gran impacto en las ventas. Es el caso de, por ejemplo, ropa de baja o mediana calidad.

Para este experimento, como se ha explicado anteriormente, se asume para simplificar la simulación, que la cantidad de unidades producidas es la vendida. Sin embargo, para reflejar de cierta manera el impacto que puede tener el precio en la variable unidades, se van a establecer algunas penalizaciones por un precio muy alto o muy bajo. La definición de muy alto o muy bajo se basará en el margen sobre el coste que la empresa se lleva. El margen estándar de una *start-up* sobre su producto, cuando se trata de una empresa que es un *marketplace* suele estar entre el 7% y el 8%[16]. A pesar de que en ciertas áreas los márgenes tienden a ser sustancialmente más altos, como en las industrias de SaaS y tecnología, que suelen ser entre un 75% y un 80%, lo habitual en el resto de las empresas es lo mencionado anteriormente. Por tanto, con efectos de una simulación, se empleará el primero.

De tal manera, las siguientes penalizaciones se aplicarán sobre los beneficios:

- Si precio $(z) > 1{,}10 * coste\ unitario$, los beneficios se penalizarán en un 5%
- Si precio $(z) > 1{,}15 * coste\ unitario$, los beneficios se penalizarán en un 10%
- Si precio $(z) < 1{,}05 * coste\ unitario$, los beneficios se penalizarán en un 5%
- Si precio $(z) < 1{,}03 * coste\ unitario$, los beneficios se penalizarán en un 10%

ii. <u>Rango de oscilación del precio</u>

Partiendo de la base de que el coste estándar, es decir, sin descuentos, de producción por unidad es de 1.000€ (como se verá en el apartado siguiente), por tanto, se ajustarán los rangos del precio entre 1.000€, que es lo mínimo para cubrir el coste variable y 1.200€, un margen del 20% sobre el coste variable.

$$Bounds[z] = [1.000, 1.200]$$

b) Unidades de producción (x)

i. <u>Comportamiento de la variable</u>

En una función simple de beneficios, estos de determinan como ingresos menos costes. Sin embargo, algunos de esos costes se tratan de costes variables, que dependen de las unidades producidas, y otros se tratan de costes fijos, que se determinan relativamente independiente de

[16] Según la socia fundadora de K Fund

las unidades producidas. Los ingresos, por su parte, dependen de las unidades producidas y del precio. De tal manera, las ecuaciones interiores se expresan de la siguiente manera:

$$Ingresos(x, z) = unidades(x) * precio(z)$$
$$Costes\ variables(x) = unidades(x) * coste\ unitario$$

Como se puede observar en las ecuaciones anteriores, al haber asumido que las unidades producidas son las mismas que las unidades vendidas, los ingresos, que dependen de unidades y precio, y los costes variables, que dependen de unidades, están relacionados. Por tanto, las variables unidades, x, y precio, z, están relacionadas entre sí.

De esta manera, teoréticamente, la parte correspondiente a estas dos variables de la expresión matemática de los beneficios sería:

$$Beneficios(x, z) = x * z - coste\ variable\ (x)$$

Basándonos en la función anterior, habría que determinar cómo se calcula ese coste variable dependiente del número de unidades.

El coste variable de producción está relacionado con la compra de materiales o recursos que se necesitan para fabricar el producto en cuestión u ofrecer el servicio. ¿Qué ocurre con estos costes? Se dice de manera simplista, que el coste variable está directamente asociado a las unidades de producción. Es decir, que el coste unitario desde el punto de vista del coste variable no cambia. Sería, siguiendo esta concepción, una variable con un comportamiento lineal, cuya pendiente sería el coste unitario constante. De ahí la simple función de unidades multiplicado por coste unitario. Sin embargo, en el mundo real no se aplica de esta manera.

En el mercado se aplican los descuentos por volumen. Ello implica, que a medida que se compra mayor volumen, se realizan ciertos descuentos por volumen. De tal manera, que el coste unitario variable va disminuyendo, de manera muy sutil, a medida que se van adquiriendo más unidades. Este comportamiento de esta variable se debe tomar en consideración a la hora de elaborar la función objetivo. De cara a la misma, el coste variable opera como una secuencia geométrica. Se toma como referencia el coste del material para una única unidad y a ello se le aplica un descuento. En la mayoría de las ocasiones, se trata de un descuento progresivo, de tal manera que el coste de añadir una unidad adicional es inferior al que fue el coste de añadir la unidad adicional anterior. ¿Cómo se presenta esto? Como un sumatorio de una secuencia geométrica.

Se tomará como asunción que el coste del material y recursos para producir una unidad es de 1.000€. Se trata de un producto simulado de coste alto. De tal manera que, si sólo se quisiese producir una unidad, este sería el coste de los materiales para producirla. A medida que se van comprando más unidades, el coste de añadir una unidad adicional es inferior ya que se aplica el descuento por volumen de venta. Se tomará como asunción que el descuento que se aplica de manera progresiva en la secuencia geométrica es del 0,01% cada vez que se añade una unidad. De esta manera, ilustrativamente, el precio de la segunda unidad sería el siguiente:

$$Unidad\ 2 = 1.000 * (1 - 0.0001)^{2-1} = 999.00€$$

O expresado matemáticamente:

$$a_x = coste\ material \sin descuento * (1 - \%\ descuento)^\wedge(x - 1)$$

Como se puede observar en la ecuación anterior, el precio de la segunda unidad es menor al precio de la primera unidad, únicamente un céntimo menor, sin embargo, cuando se compran grandes volúmenes el descuento empieza a ser más significativo. El coste de comprar materiales para x unidades será el sumatorio del coste de comprar materiales para $x - 1, x - 2, x - 3\ldots$ así sucesivamente hasta sumar el coste de la primera unidad. Ello se expresa con la siguiente función de sumatorio de secuencias geométricas:

$$S_x = \frac{1.000 * (1 - (1 - 0.0001)^x)}{(1 - (1 - 0.0001))}$$

O expresado matemáticamente:

$$S_x = \frac{coste\ material \sin descuento * (1 - \%\ descuento^x)}{(1 - \%\ descuento)}$$

La anterior función se representa de la siguiente manera:

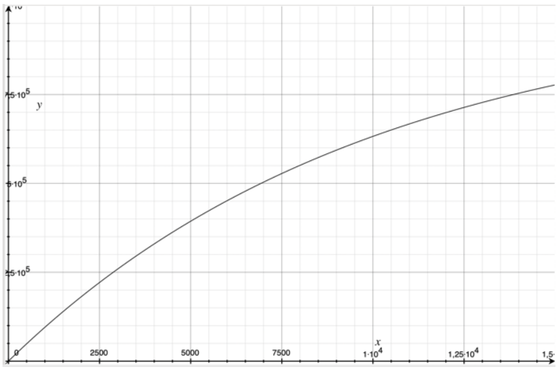

Imagen 10: función representativa de ecuación de) unidades producidas (x) vs. coste de materiales (y)

Esto implica que, al combinarlo con la función de los ingresos restándole el coste a la misma, obtendríamos el beneficio en términos de las unidades producidas.

ii. Rango de oscilación de las unidades de producción

El rango dentro del cual se buscará optimizar las unidades de producción será de entre 500 unidades y 1.900 unidades.

$$Bounds[x] = [500{,}1.900]$$

Si se producen 1.600 unidades, el coste de las mismas será de 798.265€, aproximadamente un 80% del capital disponible, mientras que, si se producen 500 unidades, el coste será de 393.621€, aproximadamente un 40% del capital total disponible de la empresa. No es factible asumir que la empresa va a gastar menos de un 20% en sueldos, costes fijos y reservas, por lo que al menos un rango. Por otro lado, tampoco resulta razonable que una *start-up*, que es una empresa que está empezando, emplee menos de un 40% de su capital disponible en su producto, que es su fuente de ingreso.

c) Número de trabajadores (v)

i. Comportamiento de la variable

El número de trabajadores es la segunda variable que será introducida en la función objetivo. En este caso, se trata de una variable con un comportamiento cuadrático negativo, es decir, $-v^2$. La función cuadrática se puede representar ilustrativamente de la siguiente manera:

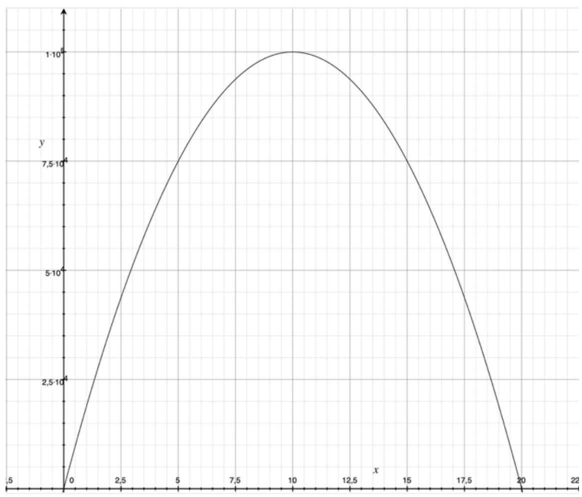

Imagen 11: función cuadrática ilustrativa $y = -1.000(v - 10)^2 - 100.000$

El número de trabajadores que tiene una empresa está directamente relacionado a sus gastos, pero a su vez con sus beneficios. Se trata de una variable, que no está directamente relacionada con las unidades de producción ni con los costes fijos que potencialmente puede tener la empresa. Es un factor distinto a los anteriores. El número de empleados incluye aquellos que se encuentran en la administración de la empresa, como aquellos que se encuentran en la producción. En parte relacionado con lo visto con la variable de producción, el número de empleados es algo que debe estar en su punto óptimo.

Los trabajadores traen beneficios, ya que aportan su saber hacer, sus contactos y sus clientes a la empresa. No se trata de un beneficio tan cuantificable como los ingresos por ventas, pero se trata de un beneficio que se debe considerar. Al principio, con muy pocos trabajadores, los beneficios son muy bajos. El tener pocos trabajadores implica que la posibilidad de producción es limitada. Para poder obtener el proceso de producción más eficiente y, en consecuencia, tener los máximos beneficios, se debe tener un equipo en el que haya una gran variedad de conocimientos. Lo ideal, es obtener un equipo de profesionales con amplia experiencia en distintos campos y juntar esas experiencias para una producción óptima. Así es como crecen las grandes *start-ups*. Ese número óptimo de trabajadores es el punto máximo de la función de la variable. Una vez se supera ese número, los beneficios empiezan a descender. Se tiene más trabajadores de los que se debe tener y el trabajador adicional cuesta más de lo que aporta.

Para efectos de la función objetivo de la presente simulación, la función cuadrática tomará la forma de:

$$y = -1.000(v - 10)^2 - 100.000$$

Al tratarse de un experimento, se han adoptado números de trabajadores simulados, por lo que, si se tratase de una aplicación real, los rangos de trabajadores dependerían del tipo de empresa. Para este supuesto, se ha adoptado una función cuyo máximo se encuentra en 10 trabajadores, y corresponde con 100.000€ en beneficios. Si la variable trabajadores implicase únicamente la obtención de beneficios, el optimizarla sería muy sencillo, pero no es el caso. También se debe pagar el sueldo de los trabajadores, por lo que se incluye en la expresión la substracción de los mismos.

De cara a la expresión de sueldos, se tomará la media de salarios en España como referente y se asumirá que se distribuyen las posiciones de alto cargo o personal altamente cualificado con las posiciones de bajo rango de tal manera que se asume una que los sueldos de manera agregativa entre el número de trabajadores es equivalente al salario medio. Según el Instituto Nacional de Estadística, el salario medio en España para el año 2021 fue de aproximadamente 25.000€ al año[17]. Por tanto, se multiplicaría este sueldo a pagar por el número de trabajadores y se realizaría una sustracción del coste. La expresión final para esta variable sería la siguiente:

$$y = -1.000(v - 10)^2 - 100.000 - 25.000 * v$$

[17] Instituto Nacional de Estadística, "Decil de salarios del empleo principal", *INE*, 2022 (disponible en: https://www.ine.es/prensa/epa_2021_d.pdf)

ii. Rango de oscilación del número de trabajadores

Expuesto lo anterior, los rangos de la variable u, número de empleados, de acuerdo con la función de beneficios por parte de los trabajadores, se situarán entre 5 y 20.

$$Bounds[u] = [5,20]$$

Desde el punto de vista de los beneficios, la primera parte que compone la ecuación de esta variable, situamos el rango inferior en 5. Se puede entender que, a pesar de que la función de un beneficio a la empresa por un único trabajador se requiere al menos de cinco para poder operar la empresa, para tener el conocimiento requerido y la mano de obra. Por otro lado, tener más de 20 trabajadores no parece razonable al ser una *start-up*, una empresa que recién está empezando y que requiere de un equipo que tenga el conocimiento necesario, pero no excederse en personal.

Desde el punto de vista de costes, la segunda parte que compone la ecuación de esta variable, el máximo coste de sueldos que tendrá la empresa será 500.000 euros, que ya de por sí es un coste que supone el 50% del capital disponible. Emplear más del 50% no parece razonable considerando que se trata de una *start-up* que debe emplear eficazmente sus limitados recursos.

d) Gastos fijos ($\boldsymbol{w}$)

i. Comportamiento de la variable

Aparte de los ingresos y los costes variables, los cuáles computan el margen de contribución, además de los costes de los sueldos de los trabajadores, de cara a los beneficios de la empresa se deben considerar también los gastos fijos. Son aquellos que son independientes de las unidades que se producen. Estos gastos normalmente incluyen el alquiler o alquileres, además del agua, la luz y demás gastos administrativos. Para efectos de este experimento, se considerarán todos los anteriores como fijos y se determinará cuál es el coste óptimo que se debe permitir gastar la empresa en esta partida para optimizar los beneficios.

Esta es una variable que se representa de manera sencilla. A medida que aumentan los costes, disminuyen los beneficios, y viceversa. Por tanto, los costes fijos en relación con los beneficios se representan matemáticamente con una ecuación lineal. Únicamente, se debe restar al agregado de las ecuaciones anteriores el coste fijo, de tal manera que en la expresión matemática de la función objetivo esté la expresión $-w$.

ii. Rango de oscilación de los gastos fijos

El rango que se podrá gastar en el concepto de gasto fijo será de 9.600 a 60.000.

$$Bounds[v] = [9.600, 60.000]$$

No resulta factible que los gastos fijos de la empresa impliquen menos de 800€ al mes. Se debe pagar el alquiler, además del agua, la luz, y demás gastos del loca. Por tanto, este será el límite inferior de la variable. Por otro lado, no parece razonable que la empresa, siendo una *start-up* con un máximo de 20 trabajadores, alquile una oficina de más de 5.000€ al mes. Estamos ante un supuesto simulado, sin embargo, en un caso real, se debería tomar en consideración particularmente si se trata de un local comercial, en el cual la localización es de gran importancia.

e) Reservas (u)

i. Comportamiento de la variable

Las reservas son una cantidad económica que se alade a la línea base de costes del proyecto para obtener un presupuesto. Es decir, la cantidad que se reserva por si surge algún imprevisto.[18] El objetivo de cualquier empresa es el de obtener beneficios, pero existe una parte de los mismos que se retiene con el objetivo de elevar la solvencia.

[18] Huerta, J. M., "Reservas de contingencia y reservas de gestión", *Gestión en TI*, 2014 (disponible en: https://josehuerta.es/gestion/proyectos/costes/reservas-de-contingencia-y-reservas-de-gestion)

¿Qué tipo de reservas existen? Los dos principales tipos de reservas son la de contingencia y la de gestión. La primera se trata de una parte del presupuesto que está destinado a cubrir los eventos o riesgos que se prevén y sale de la incertidumbre que se tiene con el proyecto en particular. La segunda, se trata de una parte del presupuesto destinada a cubrir los eventos inesperados y se suele calcular sobre toda la empresa, no solamente sobre el proyecto en particular. ¿Cómo se cuantifican estas reservas? Se debe tomar en consideración el margen de confianza sobre la estimación del presupuesto de la empresa y, a su vez, la previsión de riesgos.

La previsión de riesgos es un tema esencial en la planificación financiera de una empresa, ya que define su supervivencia a largo plazo. Sin embargo, se trata de un análisis muy complejo. Para efectos de la simulación de este trabajo, la variable se comportará de una manera lineal, al igual que el coste fijo, $-u$.

ii. Rango de oscilación de las reservas

Para las reservas a retener por parte de la empresa, el rango oscilará entre 0 y 100.000.

$$Bounds[u] = [0{,}100.000]$$

Debido a que las reservas de carácter obligatorio, es decir, las legales, se retienen de los beneficios obtenidos por la empresa en un año. Se retiene, por lo menos, un 10% del beneficio anual hasta alcanzar un 20% del capital social. A pesar de que no se le obligaría a la empresa a retener una parte del capital disponible, se puede optar por ello. Esto proporcionaría más seguridad a la empresa al tener más solvencia.

A pesar de que no conviene retener un 20% del capital, ya que esta cifra se supone que se alcanza pasados unos años, podría ser conveniente retener el porcentaje que legalmente le retendrían a los beneficios. Por ello, el límite alto de las reservas estará situado en el 10% del millón de euros disponible. Por otro lado, el límite inferior estará en 0. No es obligatorio que la empresa retenga nada.

3. PLANTEAMIENTO MATEMÁTICO DE FUNCIÓN OBJETIVO

El planteamiento de la función objetivo es esencial de cara a testear la hipótesis de la optimización bayesiana. De cara a ver si verdaderamente este método de optimización puede funcionar. La función que se va a optimizar es la función objetivo. Como se ha explicado anteriormente, en un supuesto cotidiano en el que se aplicaría la optimización bayesiana, la función objetivo es desconocida. Sin embargo, para efectos de este experimento, se programará dicha función para luego ser testeada. Por ello, se debe analizar cómo se programa dicha función.

La función objetivo es el conjunto de todas las expresiones matemáticas explicadas en los apartados anteriores. Por tanto, la función beneficios dependerá de las variables unidades, precio, empleados, coste fijo y reservas, representadas con x, z, v, w, u. Si combinamos las expresiones explicadas, la expresión matemática final será la siguiente:

$$\begin{aligned} &Beneficios(x, z, v, w, u) \\ &\qquad = x * z - 1.000 * \frac{(1 - (1 - 0.001)^x)}{1 - (1 - 0.001)} - 1.000(v - 10)^2 - 100.000 \\ &\qquad - 25.000 * v \\ &\qquad\qquad -w - u \end{aligned}$$

Tras obtener la función de beneficios, se deben incorporar las penalizaciones por precio:

- Si precio $(z) > 1{,}10 * coste\ unitario$, los beneficios se penalizarán en un 5%
- Si precio $(z) > 1{,}15 * coste\ unitario$, los beneficios se penalizarán en un 10%
- Si precio $(z) < 1{,}05 * coste\ unitario$, los beneficios se penalizarán en un 5%
- Si precio $(z) < 1{,}03 * coste\ unitario$, los beneficios se penalizarán en un 10%

Como se mencionó al principio de este apartado, se parte de la base de que la empresa tiene un capital disponible de un millón de euros. Los rangos de oscilación de las variables han sido establecidos con esta restricción en mente. Finalmente, lo que deseamos obtener es el valor óptimo de cada una de las variables, pero debemos restringir nosotros el entorno de búsqueda, estableciendo los máximos y mínimos de cada variable. Dicho esto, es posible que el resultado final combine las variable de tal manera que el gasto, entendido como los costes más las reservas. Si ese es el caso, se debe aplicar una penalización al exceso de presupuesto.

4. COMPARACIÓN DE RESULTADOS[19]

Tras realizar el proceso de optimización bayesiana sobre la función matemática descrita en el apartado anterior, los resultados obtenidos son excelentes. La comparativa se ha realizado con dos metodologías: la optimización bayesiana y la búsqueda aleatoria. A continuación, se comentarán las tres veces que se ha ejecutado el código, variando el número de experimentos realizados en cada una.

Cuando se realiza un experimento para testear una posible aproximación a un problema, el experimento debe ser lo más estadísticamente fiable posible. Esto se consigue realizando el mayor número de experimentos posibles, ya que para cada experimento se coge una aleatoriedad distinta. El aumentar el número de experimentos reduce la posibilidad de que un resultado positivo de cara al funcionamiento de la aproximación, en este caso, la posibilidad de que el hecho de que la optimización bayesiana funcione sustancialmente mejor que la búsqueda aleatoria, se deba a la suerte.

Por otro lado, el número de iteraciones indica la cantidad de veces que se repite el método gaussiano y la optimización bayesiana en sí dentro de un experimento. Para simplificar la comparativa de resultados, se ha ajustado el número de iteraciones a 10 en los dos primeros escenarios. De tal manera, se puede observar lo que ocurre al modificar el número de experimentos. Por último, en el tercer escenario, se aumenta el número de iteraciones a 30. Esto permite ver cómo evoluciona la BO cuando se realizan más procesos iterativos de búsqueda gaussiana.

[19] El código del experimento se encuentra disponible en: https://colab.research.google.com/drive/1UOhtxRuUC_eM5Mg-QZsXthEPfd-EEHti?usp=sharing

a) 20 experimentos y 10 iteraciones

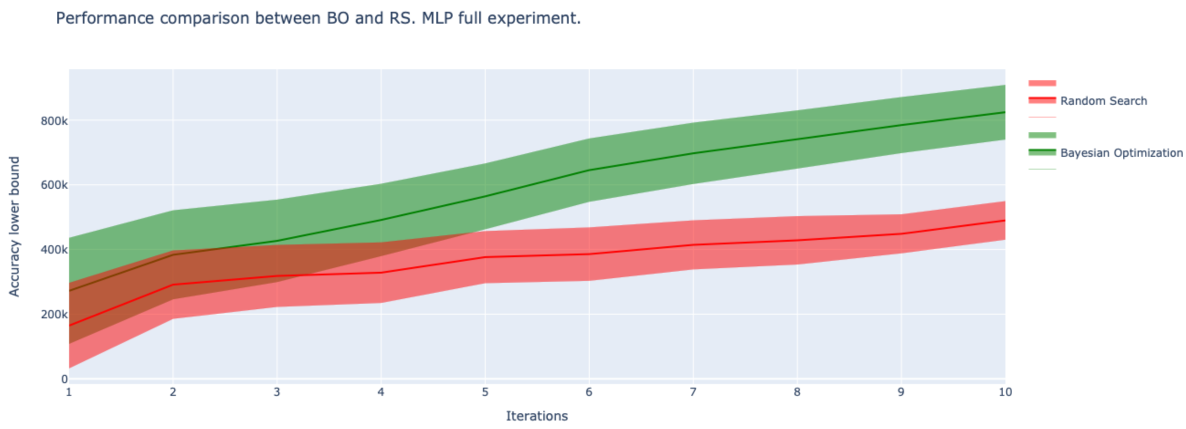

Imagen 12: comparación de rendimiento de optimización bayesiana vs. búsqueda aleatoria con 20 experimentos y 10 iteraciones

En la imagen anterior se puede observar el resultado de ejecutar el código de optimización bayesiana con una programación para que realice 20 experimentos con 10 iteraciones cada uno.

La media obtenida de beneficios en cada uno de esos experimentos es de 824,587€ con optimización bayesiana versus 489,978€ con un proceso de búsqueda aleatoria. Se puede observar, que la desviación típica, reflejada en la imagen con una sombra verde para la optimización bayesiana y una sombra roja para la búsqueda aleatoria, es superior para la primera que la segunda. Ello implica, que el nivel de variación de los resultados ha sido mayor en el proceso de BO que en el de búsqueda aleatoria. Sin embargo, también se puede observar que, a partir de la quinta iteración, el límite inferior de BO supera al límite superior de la búsqueda aleatoria, por lo que resulta ser, sin ninguna duda, el procedimiento óptimo.

El mejor resultado obtenido en esta ejecución del experimento es de 998,355.35€. Este resultado se da en el experimento número 19 y en la octava iteración, cuando las variables adoptan los siguientes valores:

- Unidades: 1,900
- Precio: 1,200

- Empleados: 5
- Coste fijo: 11,759
- Reservas: 0

Observamos de los resultados anteriores, que las unidades de producción se llevan al máximo, al igual que el precio, que los empleados se llevan al mínimo, que el coste fijo está en una posición cercana al mínimo y que las reservas son inexistentes. Estos resultados hacen mucho sentido dentro de una estrategia empresarial que considera las asunciones vistas y que únicamente toma en consideración estas medidas. La empresa va a poner el precio más alto posible (siempre que se asuma que no afecta a la demanda, asunción que no se puede materializar en la práctica), y va a vender el máximo número de unidades posible. Por otro lado, el coste de los empleados es elevado, por lo que la empresa intentará mantenerlos al mínimo. El coste fijo es prácticamente el límite inferior del mismo a su vez, es decir, busca minimizar el mismo, y las reservas son 0, manteniendo la voluntariedad de las mismas que, siempre que el objetivo sea obtener los máximos beneficios, dicha voluntariedad será no retener reservas.

b) 100 experimentos y 10 iteraciones

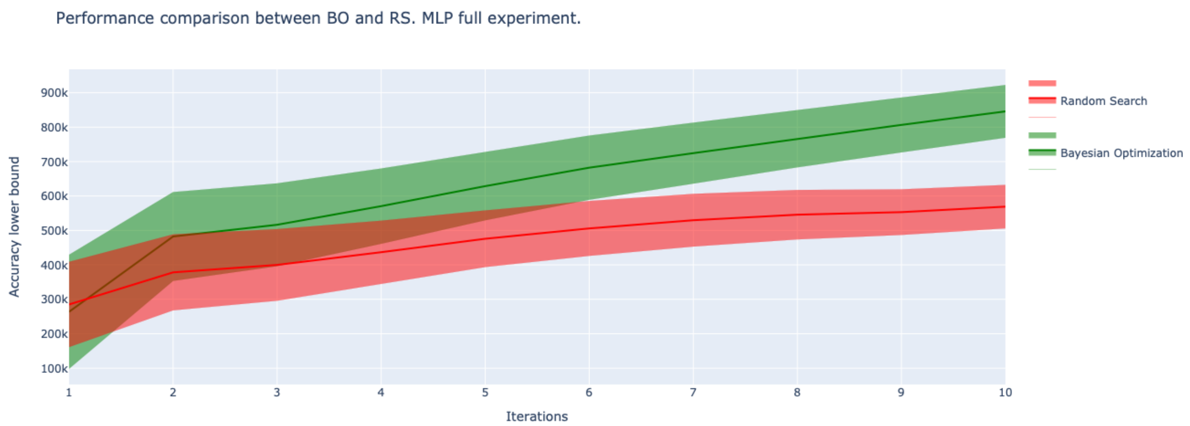

Imagen 13: Comparación de rendimiento de optimización bayesiana vs. búsqueda aleatoria con 100 experimentos y 10 iteraciones.

En la imagen anterior podemos observar como la desviación típica en cada iteración aumenta versus lo enseñado en experimentos más pequeños. Esto se debe efectivamente a que, al realizar 100 experimentos, hay 100 resultados. Por tanto, la desviación típica de esos 100 resultados proporciona una imagen más real dentro de la investigación. Como se mencionó anteriormente, el incrementar el número de experimentos incrementa la aleatoriedad de la prueba y, por tanto, los resultados se encuentran más dispersos.

De igual manera, no cabe duda alguna de que la optimización bayesiana resulta el mejor método para encontrar la estrategia óptima empresarial. Se puede observar que la búsqueda aleatoria ha encontrado un máximo de media de 569,095€, mientras que la optimización bayesiana, en su media, genera 845,744€.

Observamos con estos resultados que, si los comparamos con el apartado anterior de 20 experimentos, la optimización bayesiana da un número un poco superior, pero no del todo distinto. Esto indica que en los 20 experimentos no hubo de cara a los resultados un componente de suerte. Sin embargo, el tener 100 experimentos aporta credibilidad y fiabilidad en ese muy parecido número. Asimismo, volvemos a observar cómo la optimización bayesiana supera con creces las expectativas de una búsqueda aleatoria.

El mejor resultado obtenido en esta ejecución del experimento es de 999,299.38€. Este resultado se da en el experimento número 48 y en la décima iteración, cuando las variables adoptan los siguientes valores:

- Unidades: 1,900
- Precio: 1,200
- Empleados: 5
- Coste fijo: 9,768.4
- Reservas: 886,9

Podemos observar de estos resultados que, al igual que con la ejecución de 20 experimentos, las unidades se llevan al máximo y el precio también, mientras que los empleados al mínimo. En esta ocasión, el coste fijo que proporciona los beneficios más altos encontrados es más próximo al límite inferior. Por otro lado, en esta ocasión, la optimización bayesiana ha optado por aportar 887€ a reservas.

c) 25 experimentos y 30 iteraciones

Imagen 14: Comparación de rendimiento de optimización bayesiana vs. búsqueda aleatoria con 25 experimentos y 30 iteraciones.

En la imagen anterior podemos observar como la desviación típica de la optimización bayesiana se reduce sustancialmente. Esto se debe a un aumento del número de iteraciones. En cada iteración, la computadora realiza una nueva computación gaussiana y reajusta el punto de búsqueda. A medida que incrementamos las iteraciones, se realizan más búsquedas y se vuelven cada vez más precisas. Esta es la razón por la que la precisión de la respuesta en la iteración número 30 es muy superior a la precisión de la respuesta en la iteración 10.

Con este ejemplo, para concluir, logramos ver cómo, nuevamente, la optimización bayesiana resulta la metodología óptima. A medida que realizamos más iteraciones se vuelve más clara la eficacia de la metodología. Se puede observar que la búsqueda aleatoria ha encontrado un máximo de media de 644,902€, mientras que la optimización bayesiana, en su media, genera 1,000,782€. Volvemos a observar cómo la optimización bayesiana supera con creces las expectativas de una búsqueda aleatoria.

El mejor resultado obtenido en esta ejecución del experimento es de 1,000,201.65€. Este resultado se da en el primer experimento y en la iteración número 18, cuando las variables adoptan los siguientes valores:

- Unidades: 1,900
- Precio: 1,200

- Empleados: 5
- Coste fijo: 9,600
- Reservas: 0

Podemos observar de estos resultados que las unidades se llevan al máximo y el precio también, mientras que los empleados, el coste fijo y las reservas al mínimo. Finalmente, concluimos que el mejor valor de beneficios obtenido en todas los experimentos realizados es de 1,000,201.65€. Siendo este superior a cualquiera obtenido en búsqueda aleatoria.

CAPÍTULO VIII: CONSIDERACIONES ADICIONALES

A lo largo de este trabajo, se ha desarrollado una metodología para optimizar los beneficios de una *start-up.* Sin embargo, como se ha mencionado durante la exposición de la investigación, el trabajo parte de ciertas asunciones para simplificar la aplicación de la optimización bayesiana. Una de las principales asunciones siendo que los beneficios únicamente dependen de las variables explicadas. En la práctica, esto no sucede de tal manera, y los beneficios de una empresa dependen de cientos de variables que se interrelacionan entre sí. Por ello, si se quisiese generar un programa extenso para aplicación real, se deberían incluir, entre otras, algunas de las variables que se explicarán a continuación.

1. CAPITAL DISPONIBLE

Uno de los temas de crucial importancia al desarrollar una estrategia de negocio es analizar de cuánto capital se dispone. En este caso, se ha explicado de donde procede dicha financiación y en qué rangos se mueve. Para tomarla en consideración, los límites de las variables se han ajustado pensando en este número, en este capital disponible. Sin embargo, en la práctica, si se fuese a emplear esta metodología para un modelo real de empresa, se debería limitar de manera expresa.

Para aplicar una limitación de presupuesto, una forma de hacerlo sería con un modelo probabilístico jerárquico, ya que dentro de un proceso gaussiano no se podría modelar. El modelo probabilístico jerárquico implica que los niveles de jerarquía son una variable, y el valor que se obtiene para una variable modifica el rango de otra variable, es decir, los rangos de las variables están interconectados. De tal manera, se programaría un modelo que no sobrepase el presupuesto de costes establecido.

2. UNIDADES DE PRODUCCIÓN COMO EXPRESIÓN INDEPENDIENTE

Además de considerar las unidades de producción dentro de las expresiones matemáticas para determinar el coste del material y los ingresos, esta variable también tiene un impacto por sí sola.

Esta variable sigue el comportamiento de una función cúbica, x^3. Las función cúbicas se puede representar ilustrativamente de la siguiente manera:

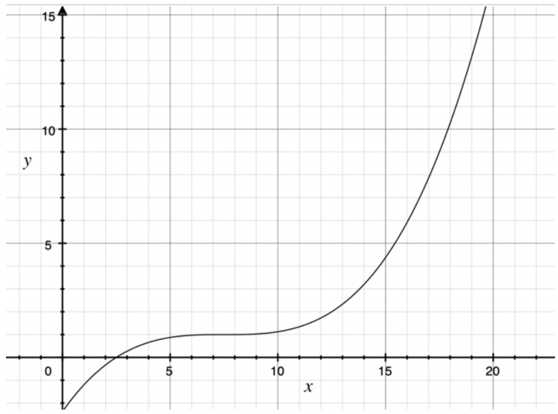

Imagen 15: función cúbica ilustrativa $y = \left(\frac{x}{5} - 1.5\right)^3 + 1$

Se dice que la función de producción en relación con los beneficios es una función cúbica. ¿A qué se debe esto? En primer lugar, como viene ilustrado en el gráfico expuesto, las primeras unidades de producción generan un resultado negativo. La empresa empieza a tener beneficios a partir del número de unidades que determina el punto de equilibrio. Esto implica el número de unidades que se deben producir para cubrir los costes. A medida que va creciendo la producción, se empiezan a obtener algunos beneficios. Sin embargo, existe un periodo de ajuste desde que se empiezan obtener beneficios hasta que estos crecen a una velocidad más rápida. Esto se da porque cuando la empresa empieza a tener beneficios, empieza a reinvertirlos para conseguir un proceso de producción más eficiente – traducido a menos costoso. Por tanto, a medida que este proceso va ocurriendo, los beneficios van creciendo. Una vez se ha llegado a un método óptimo de producción, los beneficios crecen de forma prácticamente lineal en correspondencia con las unidades de producción. A su vez, la producción a gran escala hace que los costes de disminuyan. Esto se debe principalmente a que los costes fijos de producción se distribuyen entre más unidades, por lo que el coste unitario disminuye. A este efecto lo denominamos economía de escala.

Por último, en cuanto a unidades de producción, también se debería de considerar la capacidad que tiene la empresa para producir, u ofertar, cierto número de unidades. Para ello, se debería considerar la maquinaria necesitada y se debería también relacionar con el número de trabajadores, de tal manera que quede reflejado el número de unidades que se producen por máquina y el número de unidades que produce cada trabajador. Así, se trataría de una decisión de cuánta inversión dedicar a incrementar la capacidad de producción para aumentar las unidades producidas al punto óptimo de unidades.

3. OFERTA Y DEMANDA, PRECIO Y CANTIDAD

Para efectos de esta investigación, se ha asumido que el número de unidades producidas equivale al número de unidades vendidas. En la realidad, esto no refleja el comportamiento de una empresa. Uno de los objetivos de la empresa es establecer un precio que genere un beneficio lo más superior posible sin perjudicar el número de unidades que se vende. Para esto, se debe analizar la elasticidad del bien, es decir, cómo va a reaccionar la cantidad del bien al precio del mismo. Tras este estudio, se debe incluir también un estudio de la oferta y la demanda. ¿Cuánta gente quiere el producto o servicio y cuántas compañías lo ofrecen y en qué cantidades? Este estudio, en conjunto con el anterior, determina las unidades que la compañía espera vender a cierto precio. De ahí, se establece también un rango de oscilación del mismo para poder determinar las unidades que debe haber en stock y la cadena de producción necesaria para las mismas.

4. VARIACIÓN DE PRECIOS DE RECURSOS EXTERNOS

Otra consideración adicional que se debería incluir en un modelo extensivo de beneficios es el coste por recursos externos. Dentro de esta categoría, se entraría a considerar recursos como la energía, la electricidad, el agua, así como servicios de almacenamiento en nube o externos.

Por un lado, el precio de la electricidad o cualquier otro tipo de energía depende de factores totalmente externos que se sitúan en el mercado de materias primas. Depende del precio del barril de petróleo, del gas Henry Hub, del gas europeo… Resulta de gran dificultad predecir el precio de estos servicios, pero potencialmente se podría incorporar algún modelo externo de consultores especializados en estos servicios. Por otro lado, los precios de servicios externos pueden incorporarse con cierta variación. Por ejemplo, se podría incorporar una potencial variación en precio de servicios de nube, como podrían ser los de Amazon Web Services.

5. CONSIDERACIONES MULTI-OBJETIVO

a) Beneficios versus riesgo

El *trade-off* entre el riesgo y el beneficio es una de las compensaciones más importantes de cara a establecer un plan de negocios, ya que implica el riesgo que la empresa está dispuesta a tomar. De manera directa, implica el nivel de liquidez que la empresa decide mantener debido a un potencial riesgo que puede sufrir. Esta variable se ha tratado brevemente con la variable reservas, sin embargo, va más allá. Finalmente, lo que la empresa pretende determinar es cuánto invertir y cuánto guardar, es decir, hasta qué punto le compensa tomar el riesgo de invertir su capital o bien protegerse de una posible contingencia.

Dentro de una empresa se pueden encontrar diversos tipos de riesgo. Podemos hablar de riesgos de crédito, riesgos sistemáticos, riesgos de proveedores, riesgos regulatorios, etc. Se debería realizar un estudio de los mismos y encontrar el punto óptimo de combinación de los mismos.

b) Beneficios versus responsabilidad social corporativa

En la última década, la responsabilidad social corporativa ha adquirido una importancia enorme dentro del espectro empresarial. Los famosos *ESG Scores* afectan a las empresas desde un punto de vista de imagen, pero también desde un punto de vista de negocio. Uno de los primeros sectores en incorporar este concepto ha sido el financiero. Muchos fondos de inversión o fondos de capital riesgo toman ciertos criterios dentro de este espectro como claves para poder realizar una inversión. En muchos casos, algunas empresas se quedan sin financiamiento por no cumplir estos requisitos. Por tanto, resulta cada vez más de vital importancia tomarlos en consideración y no centrarse únicamente en incrementar beneficios, sino también a la par incrementar la valorización de este parámetro empresarial. Más allá del sector financiero, cada vez más colaboradores y proveedores de servicios optan por no trabajar con empresas *ESG*. Definitivamente, no se puede ignorar este nuevo movimiento.

c) Beneficios versus visión a largo plazo

Otro de los dilemas a tomar en consideración es cuánto priorizar los beneficios ahora versus los beneficios en el futuro. Si la empresa invierte en investigación y desarrollo, esa parte de la financiación que invierte perjudica a los beneficios en el presente. Sin embargo, esa investigación y desarrollo es la fuente de la cual la empresa genera nuevos productos o servicios, o bien mejora sus existentes, lo cual lleva a unos beneficios que prosperen a largo plazo. Por ello, se debe pensar en qué porcentaje priorizar esta investigación y si merece la pena perjudicar el beneficio actual.

CONCLUSIONES

El trabajo de investigación recogido en esta memoria ha consistido en realizar una aproximación con metodología de optimización bayesiana aplicada a un modelo de estrategia empresarial de una empresa de reciente creación o *start-up*. El objetivo de esta metodología ha sido encontrar la combinación de valores de variables que diese con el beneficio óptimo. Se puede concluir que el procedimiento ha tenido éxito y ha superado las expectativas.

En el marco teórico del trabajo se fijaron tres objetivos para esta tesis de investigación. El primer objetivo era realizar una memoria del trabajo realizado en el TFG. Se puede considerar que este objetivo se ha cumplido, pues la memoria extensa relata la investigación realizada, la metodología ejecutada y los resultados obtenidos. El segundo objetivo era realizar un experimento ilustrativo de optimización de un negocio de *start-up* con optimización bayesiana. A su vez, este objetivo también se ha cumplido con creces. El capítulo VII denominado *Experimento Ilustrativo*, detalla todo el proceso seguido para determinar el problema simulado, los componentes de la función de beneficio y cómo estos se modelan, y finalmente su optimización con un proceso bayesiano en Python. El tercer y último objetivo era el de obtener resultados del experimento ilustrativo descrito anteriormente que se puedan extrapolar a *start-ups* con variables en el modelo de negocio distintas. Este último objetivo se ha manifestado positivamente en los resultados de los experimentos, al claramente quedar demostrada la superioridad de cara a la obtención de resultados de un proceso de optimización bayesiana sobre un proceso de búsqueda aleatoria. Con ello, se puede concluir que la aproximación al problema es correcta y funciona.

Este trabajo ha servido para sentar una base sobre la aplicación de la optimización bayesiana a una estrategia de negocio de una *start-up* con el objetivo de maximizar los beneficios. Si bien el experimento ha adoptado ciertas asunciones, principalmente el que los beneficios únicamente dependan de cinco variables y el que las unidades producidas se equiparen al número de unidades vendidas, dichas asunciones han servido para poder realizar un experimento ilustrativo, como se ha denominado a lo largo del trabajo de investigación. Este experimento ilustrativo pretendía explorar la utilización de esta metodología. Habiendo sustentado que efectivamente, la metodología se puede aplicar, este trabajo se podría llevar a la práctica con una empresa real.

El nivel de complejidad que presenta el mundo real supera la de un experimento ilustrativo. Por ello, cuando esta metodología se aplique en una empresa real, el número de variables será superior a cinco, la interrelación entre ellas será sustancialmente más compleja, y se deberá añadir factores como el capital disponible. A su vez, es interesante observar las posibilidades que se generan en el mundo de las consideraciones multi-objetivo. Estos, como serían el de beneficio versus riesgo, responsabilidad social corporativa o visión a largo plazo, entre otros, dan paso a un planteamiento complejo, pero cada vez más importante, en el mundo empresarial.

Concluyo que la optimización bayesiana resulta ser una herramienta de gran utilidad para empresas de reciente creación. Que su utilización para resolver un problema de beneficios puede ser tremendamente provechosa. En definitiva, es una metodología que se ha comprobado que puede considerar gran variedad de factores, y que puede considerar aún más para reflejar la complejidad del mundo comercial de hoy en día. Resulta de gran importancia hacer uso de las nuevas tecnologías que se continúan desarrollando cada día y promover el conocimiento de las múltiples herramientas que están disponibles al público, acercándolas al usuario medio, rompiendo aquellas barreras ficticias que parecen alejarlas por su aparente complejidad.

BIBLIOGRAFÍA

Antonoglou, I., de Freitas, N., Huang, A., Silver, D., Schrittwieser, J., Yutian, C. y Wang, Z., "Bayesian Optimization in AlphaGo", *DeepMind London,* 2018 (disponible en https://arxiv.org/pdf/1812.06855.pdf)

Checa, A., "Historia del Go", *Club Go Madrid,* 2007 (disponible en https://www.clubgomadrid.org/go.php)

DeepMind, "About Us", AlphaGo (disponible en https://www.deepmind.com/about)

DLA Piper, "Friends and family round vs. ángel round", *Accelerate* (disponible en: https://www.dlapiperaccelerate.com/knowledge/2018/friends-and-family-round-vs-angel-round.html)

Ekamperi, S., "Acquisition functions in Bayesian Optimization", *Let´s talk about science*, 2021 (disponible en https://ekamperi.github.io/machine%20learning/2021/06/11/acquisition-functions.html)

Garnett, R., *Bayesian Optimization*, Cambridge University Press, Cambridge, 2023.

Garrido Merchán, E. C., "Advanced Methods for Bayesian Optimization in Complex Scenarios", *Escuela Politécnica Superior Computer Science Department,* Universidad Autónoma de Madrid, 2021 (disponible en https://repositorio.uam.es/handle/10486/699441)

González, J., "Introduction to Bayesian Optimization", *Gaussian process summer school*, Sheffield University, 2017 (disponible en http://gpss.cc/gpss17/slides/gpss_bayesopt2017.pdf).

Gonzalvez, J., Lezmi, E., Roncalli, T., Xu, J., "Financial Applications of Gaussian Processes and Bayesian Optimization", Amundi Asset Management, Paris, 2019 (disponible en: http://thierry-roncalli.com/download/Bayesian_Optimization.pdf)

Huerta, J. M., “Reservas de contingencia y reservas de gestión”, *Gestión en TI*, 2014 (disponible en: https://josehuerta.es/gestion/proyectos/costes/reservas-de-contingencia-y-reservas-de-gestion).

Instituto Nacional de Estadística, “Decil de salarios del empleo principal”, *INE*, 2022 (disponible en: https://www.ine.es/prensa/epa_2021_d.pdf)

Lei, B., Kirk, T.Q., Bhattacharya, A. et al. “Bayesian optimization with adaptive surrogate models for automated experimental design”, *npj Comput Mater,* vol. 7, n. 194, 2021 (disponible en https://doi.org/10.1038/s41524-021-00662-x)

Meta, “Introduction”, *BoTorch*, 2014 (disponible en https://botorch.org/docs/introduction.html)

Meta, “Torch.unsqueeze”, *PyTorch*, 2022 (disponible en https://pytorch.org/docs/stable/generated/torch.unsqueeze.html)

Rasmussen, C. E., *Gaussian Processes for Machine Learning*, MIT Press, Massachussets, 2005.

Snoek, J., Larochelle, H. & Adams, R., *Practical Bayesian Optimization of Machine Learning Algorithms*, Cornell University, Cornell, 2012.

Ventura, P., “Cap tables y rondas, ¿a qué deberías aspirar?”, *K Fund,* 2017 (disponible en https://blog.kfund.vc/cap-tables-y-rondas-a-qué-deber%C3%ADas-aspirar-6bdc68e6f5ea)

Zakharyan, Elizabet, “Estudio de técnicas de optimización bayesiana aplicadas a operaciones de pintura con manipuladores robóticos”, Universidad Politécnica de Valencia, 2021 (disponible en: https://riunet.upv.es/bitstream/handle/10251/174640/Zakharyan%20-%20Estudio%20de%20tecnicas%20de%20optimizacion%20bayesiana%20aplicadas%20a%20operaciones%20de%20pintura%20con%20....pdf?sequence=1)