

UNIVERSIDAD TORCUATO DI TELLA

ESCUELA DE NEGOCIOS

MBA 2017

TESIS

**“Uso de modelos de forecast de demanda
para el mercado de consumo masivo
aplicado a empresa cervecera”**

ALUMNO: JUAN PEDRO BARREIRA LANG

LUGAR: CIUDAD DE BUENOS AIRES

TUTOR: PABLO ROCCATAGLIATA

FECHA: DICIEMBRE 2021

AGRADECIMIENTOS

A mi esposa por su apoyo incondicional durante toda la cursada del MBA.

A mi familia quien me ha apoyado para cumplir con cada uno de mis logros educativos.

Por último, agradezco a mi tutor quien me ha guiado durante la etapa final del MBA.

RESUMEN

Hoy en día la industria de bebidas, al igual que cualquier industria madura, se encuentra atravesando crecimientos poco significativos año a año, por lo cual, se hace imprescindible el foco en la reducción de costos maximizando la contribución marginal de cada producto vendido.

La proyección de la demanda es una actividad que tiene gran impacto, no solo en la definición de estrategias a futuro, sino que también en los costos de la empresa. Una mala proyección puede hacer que algunos productos dejen de ser rentables al elevar los costos asociados a esta.

A lo largo de la presente tesis se repasaron las bases de la preparación de un forecast estadístico. Partiendo desde los modelos matemáticos más simples para el análisis de series de tiempo, entendiendo cuál modelo aplica mejor a cada tipo de serie de tiempo. Y llegando a las herramientas disponibles más complejas que permiten hoy entender la evolución de la demanda basado en datos históricos y correlaciones con otras variables.

Adicionalmente se intentó entender cuáles son las dificultades más comunes que puede enfrentar una empresa al intentar armar un buen forecast. Y finalmente se planteó los posibles impactos a lo largo de la cadena de suministro de errores en esa proyección.

Luego se analizó el caso particular de la empresa Beers para su negocio de cervezas dentro de Argentina, se relevó el proceso de preparación del forecast así como también el proceso de preparación del plan de producción y de transporte que toman este forecast como principal input. Durante este análisis se detectaron varios puntos de dolor, identificando las principales dificultades existentes al momento de preparar el forecast así como también el impacto que se tiene de los errores del mismo.

ÍNDICE

AGRADECIMIENTOS	3
RESUMEN	4
INTRODUCCIÓN	8
MARCO TEÓRICO	10
CAPÍTULO I: Técnicas de Forecasting	10
I.I. Introducción	10
I.II. Forecast estadístico	10
I.II.I. Promedio móvil	10
I.II.II. Suavizado Exponencial	10
I.II.II.I. Simple	11
I.II.II.II. Métodos de tendencia	11
I.II.II.III. Tendencia amortiguada	12
I.II.II.IV. Suavizado exponencial Triple	12
I.II.III. ARIMA	13
I.II.IV. Modelos de regresión	13
I.II.V. Redes Neuronales	15
I.III. Machine Learning	20
I.III.I. ¿Qué es?	20
I.III.II. Consideraciones	21
CAPÍTULO II: Herramientas de Forecast	23
CAPÍTULO III: Importancia del forecast de la Demanda	27
II.I. Impacto	27
II.II. Variables que dificultan la preparación del forecast	30
MARCO EMPÍRICO	33
I. Introducción	33
II. Problema	33

III. Proceso de Forecasting	34
III.I. Planificación del forecast desde el área de Revenue Management	34
III.I.I. Preparación del forecast de ventas	34
III.I.II. Datos utilizados para el modelado	35
III.I.III. Dificultades que enfrentan	35
III.II. Planificación de la producción	36
III.II.I. Proceso de planificación de la producción	36
III.II.II. Problemas del proceso actual	37
IV. Impacto de errores en el Forecast a lo largo de la cadena de suministro	37
IV.I. Costos de reprogramación innecesarios	38
IV.II. Costos de producción	38
IV.III. Reducción de margen	39
IV.IV. Costos por pérdidas de ventas	39
IV.V. Traslados a destinos equivocados	40
IV.VI. Costo de oportunidad de capital en movimiento	40
IV.VII. Costo por derrame	40
IV.VIII. Otros	40
V. Posibles soluciones de modelos predictivos	40
V.I. Herramientas de Machine Learning Cloud	42
V.II. Herramientas de forecasting On Premise	43
V.III. Consultoría especializada	44
V.IV. Pros y Contras de cada solución	45
CONCLUSIONES	46
Bibliography	48

INTRODUCCIÓN

En la mayoría de las industrias, es preciso basarse en proyecciones a futuro para la correcta toma de decisiones, en particular, la proyección de la demanda es, sin duda, uno de los datos a considerar más relevantes. Las proyecciones a largo plazo ayudan a definir las decisiones más estratégicas de cualquier compañía mientras que las de corto plazo tienden a definir aspectos más operativos como los planes de producción con impactos directos en los costos de la compañía. “Encontrar mejoras en dichas proyecciones tendrá siempre un impacto positivo en la reducción de costos y por ende en los márgenes de la compañía.” (Oskar Blomgreen, 2016).

La industria del consumo masivo no es una excepción y la industria cervecera tampoco lo es. A su vez, algunos aspectos de esta, como los largos tiempos de preparación de los productos o la corta vida útil de los mismos tienden a hacer que estas proyecciones sean aún más relevantes a la hora de planificar la producción ya que los impactos de malas decisiones acarrearán costos aún más altos. Algunas empresas enfrentan dificultades a la hora de armar las proyecciones de la demanda, sin conseguir una precisión adecuada y por ende generando altos costos por diferencias entre la demanda real y la proyección.

Existen múltiples modelos estadísticos que proponen proyecciones desde las más simples hasta las más complejas, considerando datos históricos y variables externas que se correlacionan con la demanda. El gran avance tecnológico de las últimas décadas ofrece herramientas para abordar esta problemática de manera cada vez más sencilla para el usuario y logrando una capacidad de cálculo cada vez más compleja que permite estimar cada vez de mejor manera la demanda a futuro. Incluso la utilización de la Inteligencia Artificial aparece como tendencia para abordar problemas de forecasting.

A lo largo del siguiente trabajo se analizará la importancia del forecast de demanda entendiendo los costos asociados a una baja precisión y entendiendo que variables dificultan la misma. Luego se analizarán los modelos estadísticos recomendados para cada situación y las herramientas disponibles hoy en día.

Finalmente se analizará el caso particular de la empresa Beers entendiendo la precisión con la que se cuenta, cómo se propaga este error y que costos acarrea. Se considerarán los distintos modelos y herramientas presentadas para entender cuáles son más aconsejadas para el caso particular.

Se buscará responder las siguientes preguntas:

- ¿Qué impacto tiene para la empresa Beers la baja precisión en la proyección de la demanda?
- ¿Qué herramientas o modelos podrían aplicarse para reducir los costos asociados?

Esta Tesis es de carácter empírico no experimental del caso particular de Beers. Tiene como objetivo principal profundizar en el entendimiento del impacto de los errores de las proyecciones de la demanda sobre los costos del negocio, comprender qué herramientas existen en la actualidad y como la aplicación de las mismas pueden ayudar a la mejora de la rentabilidad del negocio.

MARCO TEÓRICO

CAPÍTULO I: Técnicas de Forecasting

I.I. Introducción

Existe una gran cantidad de técnicas y modelos estadísticos que pueden aplicarse a las proyecciones de la demanda. Saber que modelo elegir depende de muchas variables y a menudo puede sobrecargar de información a quienes tienen que implementarlos. Existen más de 70 técnicas de series temporales diferentes entre sí (Kerkkänen, 2008).

Por este motivo, en este capítulo se revisarán los conceptos básicos detrás de los modelos más conocidos y se mencionan algunos de los modelos complejos más frecuentados mencionando ventajas y desventajas de los mismos, así como también frente a qué situaciones será conveniente aplicar cada uno de ellos.

I.II. Forecast estadístico

I.II.I. Promedio móvil

El promedio móvil es uno de los métodos más simples de proyección utilizados y permite estimar valores futuros utilizando el promedio de los datos anteriores. La cantidad de datos utilizados para el promedio definirá la relevancia que se le quiera dar a los datos más antiguos de la serie.

I.II.II. Suavizado Exponencial

El uso del suavizado exponencial como método estadístico para la proyección de una serie de tiempo se basa en el modelo anterior, utilizando promedios móviles ponderados de las observaciones pasadas, con esta ponderación disminuyendo exponencialmente a medida que las observaciones se alejan en el tiempo.

Existen diversos modelos basados en el suavizado exponencial, entre los más conocidos se encuentran los siguientes:

I.II.II.I. Simple

El modelo simple es apropiado solamente para aquellas series de tiempo en las que no haya componente estacional ni de tendencia. El forecast es calculado utilizando promedios ponderados en donde el peso decrece de manera exponencial cuanto más alejada en el pasado sea la observación.

$$\hat{y}_{T+1} = \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots$$

En donde α es positivo menor a 1, es el parámetro de suavizado que define la velocidad con la que decrecen los pesos en el tiempo.

I.II.II.II. Métodos de tendencia

Dentro del suavizado exponencial surgen varios métodos que intentan resolver uno de los principales problemas del suavizado exponencial simple al incorporar en la ecuación el concepto de tendencia. Entre estos métodos se encuentra el método de tendencia lineal de Holt o también denominado suavizado exponencial doble. Se denomina de esta manera porque consta de aplicar de manera recursiva el suavizado exponencial simple dos veces.

Se rige por las siguientes ecuaciones, donde utiliza el suavizado exponencial para calcular el nivel, pero también para calcular la tendencia de la serie en un momento t

$$\text{Ecuación de Forecast: } \hat{y}_{t+h} = l_t + h \cdot b_t$$

$$\text{Ecuación de nivel: } l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1})$$

$$\text{Ecuación de tendencia: } b_t = \beta * (l_t - l_{t-1}) + (1 - \beta^*)b_{t-1}$$

Siendo α el parámetro de suavizado para el nivel y β el parámetro de suavizado para la tendencia (ambos positivos menores a 1)

Una de las principales limitaciones de este tipo de modelos es que proyectan tendencias constantes indefinidas en el futuro. Resultados empíricos indican que

este tipo de modelos tienen a sobreestimar las proyecciones de períodos de tiempo (Athanasopoulos, 2018).

I.II.II.III. Tendencia amortiguada

Con el objetivo de resolver el principal inconveniente del modelo anterior surgen los modelos de tendencia amortiguada que agregan a la ecuación un parámetro encargado de suavizar la tendencia a medida que se aleja el período proyectado de la siguiente manera:

$$\text{Ecuación de Forecast: } \hat{y}_{t+h} = l_t + (\phi + \phi^2 + \dots + \phi^h) \cdot b_t$$

$$\text{Ecuación de nivel: } l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + \phi \cdot b_{t-1})$$

$$\text{Ecuación de tendencia: } b_t = \beta * (l_t - l_{t-1}) + (1 - \beta^*) \cdot \phi \cdot b_{t-1}$$

I.II.II.IV. Suavizado exponencial Triple

También conocido como Holt and Winters método estacional y divide la ecuación de forecast en 3 ecuaciones de suavizado exponencial de manera similar al método doble, pero agregando una componente estacional con su respectivo parámetro de suavizado. Dentro del suavizado exponencial triple existen 2 variaciones que dependen de cómo se expresa la componente estacional. Existe el método aditivo y el multiplicativo, la utilización de uno u otro dependerá principalmente de si la variación estacional varía de manera proporcional con el nivel de la serie.

Suavizado exponencial triple aditivo:

$$\text{Ecuación de Forecast: } \hat{y}_{t+h} = l_t + h \cdot b_t + s_{t+h-m(k+1)}$$

$$\text{Ecuación de nivel: } l_t = \alpha(y_t + s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1})$$

$$\text{Ecuación de tendencia: } b_t = \beta * (l_t - l_{t-1}) + (1 - \beta^*)b_{t-1}$$

$$\text{Ec. de estacionalidad: } s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}$$

Suavizado exponencial triple multiplicativo:

$$\text{Ecuación de Forecast: } \hat{y}_{t+h} = (l_t + h \cdot b_t) s_{t+h-m(k+1)}$$

$$\text{Ecuación de nivel: } l_t = \alpha \frac{y_t}{s_t} + (1 - \alpha)(l_{t-1} + b_{t-1})$$

$$\text{Ecuación de tendencia: } b_t = \beta * (l_t - l_{t-1}) + (1 - \beta^*)b_{t-1}$$

$$\text{Ec. de estacionalidad: } s_t = \gamma \frac{y_t}{l_{t-1} + b_{t-1}} + (1 - \gamma)s_{t-m}$$

I.II.III. ARIMA

El Promedio Móvil Integrado Autorregresivo (ARIMA por sus siglas en inglés) junto con los modelos de suavizado exponencial son los modelos más utilizados al abordar proyecciones de series de tiempo. Mientras que los modelos de suavizado exponencial tienen como objetivo describir la componente estacional y de tendencia de una serie de tiempo, los modelos ARIMA buscan representar la serie temporal en función de rezagos de la misma y shocks aleatorios.

Las primeras dos letras “AR” refieren a los términos autorregresivos y de manera similar a como los modelos regresión buscan correlación entre una variable con predictores, se busca encontrar la relación entre la variable buscada y los valores pasados de la misma. Las últimas dos letras “MA” refieren al promedio móvil de los errores entre la ecuación de forecast y los valores reales que toma la muestra. Y la “I” refiere a que con estos modelos los valores son reemplazados con la diferencia entre en valor observado y el valor observado en el período anterior (proceso de diferenciación para convertir la serie en estacionaria)

I.II.IV. Modelos de regresión

En los modelos presentados hasta el momento es posible incluir y tomar en cuenta los valores históricos de la serie para proyectar valores futuros, pero no permite incluir información externa que puede ser relevante para determinar el comportamiento de las variables en estudio.

Un modelo de regresión busca determinar la relación existente entre 2 variables a lo largo del tiempo. En otras palabras, buscamos realizar una proyección de una serie de tiempo “y” asumiendo que tiene una relación con otra serie de tiempo “x”. Donde la variable y será la variable dependiente “y” las variables “x” las explicativas. El objetivo en forecasting es determinar la menor manera de utilizar un conjunto de variables explicativas “x” para pronosticar “y” .

El modelo más simple en este es el modelo de regresión lineal “y” asume una relación lineal entre “x” e “y” de la siguiente manera:

$$y_t = \beta_0 + \beta_1 \cdot x_t + \varepsilon$$

Una variante muy utilizada al modelo de regresión lineal es el modelo de múltiple regresión lineal, se define de esta manera al modelo de regresión lineal donde existe más de una variable explicativa, lo que define el modelo de la siguiente manera:

$$y_t = \beta_0 + \beta_1 \cdot x_{1,t} + \beta_2 \cdot x_{2,t} + \dots + \beta_k \cdot x_{k,t} + \varepsilon$$

Una medida muy utilizada para entender si este modelo aplica para establecer la relación lineal existente entre 2 variables es mediante el valor de R^2 .

$$R^2 = \frac{\sum_{t=1}^T (\hat{y}_t - \underline{y})^2}{\sum_{t=1}^T (y_t - \underline{y})^2}$$

El valor de R^2 estará cerca de 1 cuando exista una alta relación lineal entre la variable explicativa y la dependiente y será cercano a cero cuando las variables sean independientes.

Uno de los defectos de utilizar R^2 para determinar si un modelo tiene buenos predictores es que este siempre aumenta al agregar variables, lo que limita su uso para comparar modelos con distinta cantidad de parámetros. Una alternativa a utilizar en estos casos es el R^2 ajustado ya que el valor de la misma no aumenta al aumentar las variables, se expresa de la siguiente manera:

$$\underline{R}^2 = 1 - (1 - R^2) \frac{T - 1}{T - k - 1}$$

I.II.V. Redes Neuronales

Se le llama redes neuronales a una gran variedad de modelos flexibles de regresión no lineal y discriminantes, modelos de reducción de datos y sistemas dinámicos no lineales (Sarle, 1994)

Se puede considerar las redes neuronales como una red de neuronas o nodos que se organizan por capas, en donde, la entrada forma la primera capa y la salida forma la última capa y entre estas podrán existir capas intermedias. Estos elementos pueden relacionarse a través de conexiones simples o hasta muy complejas dependiendo de su uso.

El resultado obtenido a partir de la implementación de modelos de redes neuronales puede variar muy ampliamente según la arquitectura definida, los casos más básicos pueden ser similares o incluso idénticos a los modelos estadísticos mencionados previamente. La cantidad de neuronas y de capas, así como las funciones de activación y de propagación se definen previamente. Muchas veces se suele comenzar operando con una sola capa intermedia y funciones lineales.

A modo de ejemplo puede considerarse el siguiente modelo de redes neuronales con una capa intermedia:

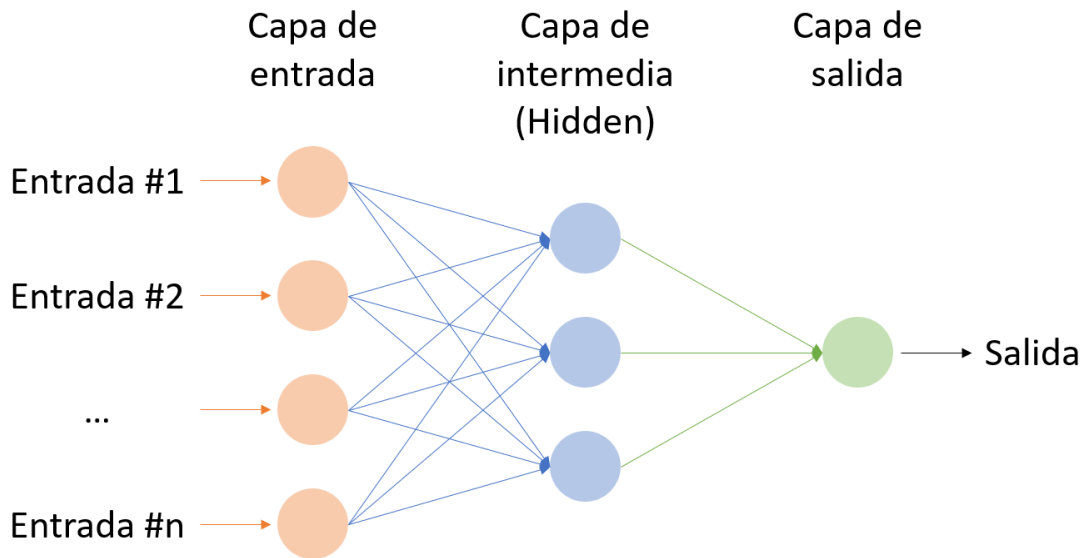


Ilustración 1: Red neuronal con una capa intermedia. Fuente: Elaboración Propia

Para la cual se define una función de propagación donde cada nodo recibe como entrada una combinación de las salidas de los nodos de la capa anterior. A esta función se le llama función de propagación. A modo de ejemplo, se puede definir una función de propagación lineal de la siguiente manera:

$$z_j = b_j + \sum_{i=1}^4 w_{i,j}x_i$$

donde w_i es el peso asignado a cada input para cada neurona de la segunda capa

Y se define también una función de activación siendo la transformación realizada por cada neurona y que define la salida de esta, utilizando la función de propagación de entrada como variable. A modo de ejemplo se puede definir un Función de activación no lineal (sigmoide) de la siguiente manera para darle la entrada a la siguiente capa, esta función reduce el efecto de valores extremos logrando un efecto robusto frente a outliers.

$$s(z) = \frac{1}{1 + e^z}$$

Las variables incluidas en estas funciones, b_j y w_i en este ejemplo, se ajustarán con iteraciones durante el aprendizaje del modelo.

Una vez que se obtiene un modelo aplicable a la necesidad planteada las funciones de activación pueden ser cambiadas al igual que las de propagación, y suele agregarse mayor cantidad de capas intermedias, lo cual le da mayor complejidad al modelo permitiendo que el mismo se adapte a situaciones más complejas, pero a su vez requiere de mayor poder de cómputo para lograr un entrenamiento adecuado.

Las limitaciones en la puesta en producción de estos modelos provienen de la infraestructura disponible para poder procesarlos y dependiendo de la aplicación, de la cantidad de datos disponibles para el correcto entrenamiento. Es común utilizar técnicas de transfer learning que permiten trabajar con modelos pre entrenados para tareas similares. De esta forma se logran acortar los tiempos de entrenamiento, así como también reducir la cantidad de observaciones necesarias para el mismo.

Los modelos de redes neuronales con varias capas intermedias suelen considerarse dentro del campo del aprendizaje profundo o Deep Learning el cual a su vez corresponde al campo del aprendizaje automático o Machine Learning.

(Harald Bauer, 2017) "Desde finales de la década de 2000, el aprendizaje profundo (Deep Learning) ha sido el enfoque más exitoso para muchas áreas donde se aplica el aprendizaje automático".

El ejemplo de red neuronal presentado previamente corresponde a un modelo simplificado dentro de una categorización de redes neuronales llamada MLP o Perceptrón Multicapa por sus siglas en inglés, siendo esta categorización la más básica dentro del mundo de Deep Learning. Al aumentar la complejidad de las relaciones entre las distintas capas surgen infinitas configuraciones de arquitectura posibles. Los tipos de redes neuronales de deep learning más relevantes, pueden ser categorizados de la siguiente manera (Lara-Benítez, 2021):

- **Redes neuronales conectadas completamente:** siendo este tipo de redes las más básicas, tal como presentado en el ejemplo anterior, su arquitectura está compuesta por 3 bloques, capa de entrada, capas escondidas y capa de salida; y las relaciones entre las neuronas son siempre hacia adelante a diferencia de las categorizaciones presentadas a continuación.
 - MLP: Multi-layer perceptron.
- **Redes neuronales recurrentes:** permite conectar cada paso con los previos para poder modelar la dependencia temporal de los datos, dándole a este tipo de redes la capacidad de soportar mejor la secuencia de datos. Dentro de la cual se encuentran las siguientes redes neuronales:
 - ERNN (Elman recurrent neural network)
 - LSTM (Long short-term memory network)
 - ESN (Echo state network)
 - GRU (Gated recurrent units network)
- **Redes convolucionales:** este tipo de redes se compone usualmente de capas convolucionales, capas agrupadas y capas completamente conectadas. Donde cada neurona se conecta solo con una región de la información de entrada y las neuronas de una misma capa comparten el mismo peso parametrizado. Estas características permiten una configuración más simple (menos parámetros de entrenamiento) logrando menores tiempos de entrenamiento, incluso siendo capaz de lograr mejores resultados. Dentro de esta categorización se encuentran las siguientes redes neuronales:
 - CNN (Convolutional neural network)
 - TCN (Temporal convolutional network)

Models	Parameters	Values
MLP	Hidden Layers	[8], [8, 16], [16, 8], [8, 16, 32],
		[32, 16, 8], [8, 16, 32, 16, 8],
		[32], [32, 64], [64, 32],
		[32, 64, 128], [128, 64, 32],
		[32, 64, 128, 64, 32]
ERNN	Layers	1, 2, 4
	Units	32, 64, 128
	Return sequence	True, False
LSTM	Layers	1, 2, 4
	Units	32, 64, 128
	Return sequence	True, False
GRU	Layers	1, 2, 4
	Units	32, 64, 128
	Return sequence	True, False
ESN	Layers	1, 2, 4
	Units	32, 64, 128
	Return sequence	True, False
CNN	Layers	1, 2, 4
	Filters	16, 32, 64
	Pool size	0, 2
TCN	Layers	1, 3
	Filters	32, 64
	Dilations	[1, 2, 4, 8], [1, 2, 4, 8, 16]
	Kernel size	3, 6
	Return sequence	True, False

Tabla 1: Grilla de parámetros comúnmente utilizados de las configuraciones de arquitectura para los 7 tipos de modelos de Deep Learning. Fuente: (Lara-Benítez, 2021)

Dentro de las redes neuronales más utilizadas presentadas anteriormente es importante hacer foco en las **LSTM**, si bien cada una puede tener usos específicos donde performen mejor sobre el resto, estas redes han demostrado obtener resultados más precisos para el forecast de series de tiempo (Lara-Benítez, 2021).

Estas redes permiten modelar dependencias temporales en largos horizontes de tiempo, pero tomando en consideración los patrones presentados en el corto plazo. Para lograr esto, LSTM agrega a los modelos presentados la capacidad de acarrear información a través de varios pasos de tiempo, previniendo la desavenencia gradual de los datos más viejos a lo largo del tiempo (Chollet, 2017). Utilizan entradas multiplicativas que previenen ser afectadas por

pequeñas alteraciones en la información. De manera similar, utiliza salidas multiplicativas para reducir el impacto de alteraciones existentes en estas memorias. Y como última característica, LSTM permite descartar información guardada a medida que estas resultan irrelevantes para el armado del forecast.

I.III. Machine Learning

Vandeput en su libro *Data Science for Supply Chain Forecasting* compara el impacto que la Inteligencia artificial tendrá las próximas décadas con el impacto que generó la electricidad a mediados del siglo 19 permitiendo que las industrias produzcan más con menos. No utilizar esta tecnología hoy en día es como utilizar velas en aquel momento compitiendo contra empresas que comienzan a sacar gran provecho de la misma (Vandeput, 2021).

“El uso de big data y el poder de cómputos masivos en la nube, la inteligencia artificial puede simultáneamente refinar cientos de modelos avanzados, mucho más allá de los que es posible con el forecasting de demanda tradicional” (Glass, 2018).

I.III.I. ¿Qué es?

Es una técnica de ciencia de datos que se centra en el aprender sobre los datos existentes con el objetivo de analizar tendencias y comportamientos futuros, esta capacidad de aprendizaje es lo hace que las aplicaciones o dispositivos sean más inteligentes (Céspedes, 2017).

A diferencia de los modelos estadísticos usuales donde se aplica un set de relaciones conocidas a un dataset, en el caso de Machine Learning se espera que el algoritmo sea quien entienda el set de datos y las relaciones dentro de este (Vandeput, 2021). Según Cassie Kozyrkov¹ al utilizar Machine Learning, “la tarea del programador pasa a ser alimentar el sistema con ejemplos y es el trabajo de los algoritmos de Machine learning definir las transformaciones a partir de estos ejemplos”

¹ <https://towardsdatascience.com/making-friends-with-machine-learning-5e28d5205a29>

El uso de tecnologías que utilizan Machine Learning para la preparación del forecast permite aplicar diferentes modelos estadísticos en paralelo y ajustar distintas variables dentro de cada uno a medida que se actualiza el modelo con nuevos datos de manera de poder comparar cual se adapta mejor a estos datos. De esta manera se obtiene un modelo que se adapte a los cambios, ajustando las variables durante el proceso, siendo capaz de incorporar nuevas dependencias de datos de manera automática o incluso de cambiar de modelos estadísticos para lograr mejores predicciones.

I.III.II. Consideraciones

Si bien se ha avanzado mucho en términos de la inteligencia artificial y el uso de machine learning para las proyecciones del forecast, para muchas empresas sigue siendo un mundo desconocido y suelen cometerse diferentes errores al momento de implementar herramientas que utilizan Machine Learning. Vanerput en Data Science for Supply Chain Forecasting (Vandeput, 2021) destaca las siguientes:

- **Calidad de la información:** la calidad del set de datos que se utiliza como input definirá finalmente la calidad de la proyección. Es por eso que es clave involucrar equipos que conozcan tanto los procesos como los productos involucrados para ser más certeros al momento de identificar eventos que puedan haber afectado la calidad de la información, así como también identificar variables que puedan afectar la proyección futura. Proveer de toda la información adecuada permitirá lograr mejores resultados en el forecast. Un error muy común es utilizar como input información de ventas en vez de información de demanda, por ejemplo.
- **Expectativas:** Tener expectativas muy altas o muy bajas es otro error que se suele cometer. Expectativas demasiado altas puede resultar fácilmente en frustraciones de la alta gerencia y puede generar una reacción reacia por parte de los planificadores de demanda. Por otro lado, expectativas muy bajas hará que este tipo de iniciativas no sean priorizadas y no se

aloque el presupuesto necesario para demostrar el alto retorno que pudiera tener.

- Poca alineación con el proceso: Antes de comenzar a trabajar en el forecast es importante tener claro qué es lo que se necesita proyectar y qué relación tiene con él/los procesos impactados. Es preciso definir el nivel correcto de desagregación de los resultados, tanto a nivel temporal como geográfico, de mercado o cualquier otro corte necesario.
- Métricas que no reflejan correctamente el objetivo del negocio: antes de trabajar sobre la mejora del forecast es preciso definir las métricas correctas para evaluar la efectividad del mismo. Un error muy común es medir el accuracy utilizando MAPE como principal métrica en vez de una métrica ponderada como el wMAPE. A diferencia del MAPE esta métrica permite ponderar los errores según la importancia de cada estimación, pudiendo ponderar por cantidad de ventas, margen, ventas totales, etc . Debemos recordar que no es lo mismo equivocarse en un 5% en el forecast de productos que dejan un margen chico vs en aquellos que tienen márgenes grandes.

CAPÍTULO II: Herramientas de Forecast

A la hora de evaluar qué herramienta de forecasting conviene utilizar, es muy importante asegurar que se adecue de manera correcta al proceso, o ajustar el proceso en caso sea necesario. Según (Basson, Kilbourn, & Walters, 2018) una herramienta de forecast ideal para poder contribuir en una mejor proyección de la demanda requiere un proceso de forecast enfocado en los siguientes puntos

- Aplicación de un modelo estadísticamente impulsado que requiera poca o nulo esfuerzo o intervención
- Uso de varias fuentes de datos distintas que provean información sobre todos los controladores primarios de la demanda
- La aplicación de un modelo, que prediga las ventas al consumidor y aprenda los patrones de orden de estos, para proveer predicciones de los envíos
- El modelo utilizado debe aprender con el correr del tiempo, ajustándose a los cambios en la demanda y en el comportamiento de las órdenes para mejorar los niveles de precisión
- Con el enfoque de forecasting mejorado, la precisión debe estar alineada y evaluada contra promedios específicos de la industria

Todos estos puntos son fundamentales para cualquier cadena de suministro que quiera desempeñarse de manera óptima y esté óptimamente soportado por la tecnología.

Existen diferentes herramientas tecnológicas disponibles útiles para la preparación del forecast y su uso dependerá de la complejidad de este, las variables se quieran incluir, los cambios de demanda a los que se enfrentan, la necesidad de precisión de este y la disponibilidad de recursos que se tengan. “La elección del software a utilizar dependerá de los métodos de forecast disponibles en cada uno, la facilidad de la entrada y gestión de datos, la calidad del forecast” (Makridakis, Wheelwright, & Hyndman, 1997). Por lo cual es necesario luego entender que ofrece y que requiere cada herramienta para poder utilizarla:

- **Capacidad de cómputo**, si bien no es estrictamente la herramienta quien limita la capacidad de procesamiento sino el hardware sobre el que se corra, en muchos casos la capacidad de procesamiento física “on premise” suele ser muy costosa o limitada, en este punto las soluciones Cloud suelen ofrecer muy altas capacidades sin tener que comprar los equipos físicos. Por lo cual es importante considerarlo previo a la elección de la herramienta, ya que elegir una herramienta en la nube permitirá posterior escalabilidad sin límites.
- **Facilidad/ practicidad de uso**, existe una muy alta variedad de opciones desde interfaces muy amigables para el usuario hasta herramientas que requieren alto conocimiento sobre distintos lenguajes de programación. Las herramientas más convencionales ofrecen ambas posibilidades permitiendo implementar buenos modelos con un bajo conocimiento de lenguajes informáticos mediante el uso de webservices o APIs. Adicionalmente una herramienta práctica tiene disponibles para fácil implementación módulos que permiten ejecutar todas las actividades necesarias dentro de estos, como ser modulo para el manejo y limpieza de datos, transformaciones, o cualquier tipo de manipulación o también visualizaciones diversas para la salida de datos de cada modelo.
- **librerías/Modelos**, un aspecto muy importante es la capacidad de cada software de trabajar con los modelos estadísticos que apliquen a la serie de tiempo. En muchos casos se suele trabajar con librerías extensas open source disponibles dentro de una comunidad enorme, esto asegura una buena flexibilidad para trabajar con los modelos que se requieran.
- **Costo**, por último, pero no menos importante, el costo del producto final. Muchas de las soluciones en la nube existentes ofrecen un servicio gratis al inicio lo que permite realizar pruebas de concepto y posteriormente suelen cobrar por capacidad de cómputo contratada a medida que se utilice, permitiendo una escalabilidad muy amplia sin ninguna inversión inicial. Por otro lado, como mencionado previamente, las soluciones on premise suelen acarrear el costo del software además del de el hardware.

Si bien con el avance tecnológico de los últimos años, las herramientas más complejas son hoy mucho menos costosas y, por ende, más accesibles para pequeñas empresas que previamente no habrían podido contar con un sistema tan sofisticado, el conocimiento sobre su uso puede continuar siendo una barrera para muchos.

Estas herramientas pueden ser separadas en las siguientes categorías según sugiere (Basson, Kilbourn, & Walters, 2018):

- Software independiente de forecast estadístico
- Paquetes de planeamiento de demanda que están integrados e incluyen forecast estadísticos, forecast de promociones, forecast del ciclo de vida de productos, e información de planning de ventas
- Paquetes de software que proveen planes de ventas y funcionalidades de gestión de relación con los clientes.
- Paquetes de software de planning colaborativo que permiten la comunicación de promociones planificadas, compras agendadas, forecast e información del punto de venta
- Conjuntos de software que incorporan todos los puntos mencionados en un software de planificación de recursos de la empresa y gestión de la cadena de suministro.

Algunos de los software para la planificación de la demanda más conocidos y disponibles hoy en el mercado son JDA Demand de JDA Software, SAP-Supply Chain Management de SAP, Forecast Management de Demand Solutions y Demantra de Oracle.

Adicionalmente a la categorización anterior es necesario mencionar que hoy en día la herramienta más utilizada para la realización de forecasts y que no es mencionada por el autor es la hoja de cálculo. Las hojas de cálculo son muy utilizadas para negocios y cuentan con infinidad de usos, muchas de estas proveen algunas herramientas para el armado del forecast a través de funciones o add-on. Sin embargo, no suelen ser la opción más recomendada dada las limitaciones que suelen tener. (Makridakis, Wheelwright, & Hyndman, 1997) “Las herramientas ofrecidas en las hojas de cálculo suelen ser insuficientes para

análisis de datos exploratorios, chequeo de supuestos del modelo o comparaciones entre varios modelos de forecasts”

Dentro de esta tesis nos enfocaremos en las herramientas especializadas en la preparación del forecast y no en las herramientas que integran varios procesos de la cadena de suministro mencionadas anteriormente. Dentro de estas herramientas es importante mencionar que, gracias a la evolución reciente en términos tecnológicos con el avance de la inteligencia artificial y el aprendizaje automático, cada vez existen alternativas más potentes, más fáciles de usar y con costos cada vez más bajos.

Existe una gran variedad de herramientas de análisis predictivo que utilizan Machine Learning para crear o implementar modelos de manera efectiva.

Como mencionado previamente existe una amplia variedad de herramientas disponibles para aplicar los distintos modelos observados previamente, incluidos los de redes neuronales de Deep Learning, incluso muchos de esos ofrecen suscripciones gratuitas en la nube y permitiendo además una posterior escalabilidad en caso de optar por modelos complejos que requiera amplia capacidad de cómputo.

Herramientas On Premise:

- Forecast Pro
- COLIBRI
- ClosePlan
- DataRobot
- Effectmanager
- FutureMargin
- Pipedrive
- Smart Demand Planner

Herramientas Cloud:

- Google ML
- Facebook Prophet

- Azure Machine Learning Studio
- Amazon Forecast
- IBM Machine Learning (o Machine Learning Watson Studio)

Es importante destacar que las herramientas de machine learning bien utilizadas suelen obtener mejores resultados que las herramientas de forecasting estadístico tradicionales. “Generalmente los modelos de machine learning le ganan a los software de forecasting por 5 a 15%. E incluso mejor accuracy puede ser obtenido a medida que más información se le incluye en el modelo” (Vandeput, 2021)

O como también comenta Glass “El modelo óptimo puede ser luego aplicado al más alto nivel de granularidad SKU-ubicación para generar un forecast que en general mejorará la precisión en más de 10%” (Glass, 2018).

CAPÍTULO III: Importancia del forecast de la Demanda

II.I. Impacto

Tal como se mencionó anteriormente, el forecast de demanda cumple un rol muy importante dentro de la operación de cada compañía en cada industria, este suele ser utilizado con varios propósitos a través de las diferentes áreas: producción, compras, logística, finanzas, marketing y ventas entre otras. Por lo cual el error de este muchas veces termina teniendo implicancias en cada una de las áreas mencionadas, generando, en algunos casos, costos asociados muy altos.

El impacto potencial de errores en el forecast ha sido revisado por varios autores, algunos de estos errores tendrán una implicancia directa y posiblemente también será más fácil dimensionarlos; por otro lado, existen impactos indirectos en donde la relación no será tan fácil de dimensionar. Los primeros son resumidos en las siguientes categorías según propone Annastiina (Kerkkänen, 2008)

- Impactos de planeamiento, se refiere a los costos asociados al exceso de trabajo de planeamiento y está principalmente afectado por la inestabilidad de la programación.
 - Costos de reprogramación innecesarios: Las inestabilidades en la programación muchas veces llevan a la necesidad de ajustes de la

programación sin anticipación para cumplir con las nuevas demandas proyectadas, reconfiguraciones innecesarias en las líneas productivas para realizar cambios de productos generan extra-costos.

- Impactos de capacidad, se refiere a los costos asociados a las pérdidas en la capacidad.
 - Costos de producción. Corresponde a la producción de productos que no se venderán
 - Uso ineficiente de la capacidad. Si una planta está utilizando su capacidad al máximo, errores en el forecast de demanda generarán pérdidas de oportunidad de producir otros productos necesarios. Este concepto tiene un impacto directo en los costos por Stock Out mencionados posteriormente y no debe duplicarse el costo asociado.
- Impactos de inventarios, se refiere a los costos asociados al manejo inadecuado de los niveles de inventario.
 - Exceso de inventarios, se refiere a los costos de mantenimiento de este exceso de inventarios, diferencia entre la demanda proyectada y la real. Estos costos incluyen los costos de almacenamiento como también de transporte.
 - Políticas de inventario equivocadas. Bajos valores de accuracy del forecast de demanda implica una mayor variabilidad de la demanda esperada y por ende será necesario mantener un mayor stock de seguridad para asegurarse un mismo cumplimiento o service level; esto aplica tanto en producto terminado como en materia prima aumentando el costo total de mantenimiento de inventarios.
 - Obsolescencia. La demanda proyectada en exceso muchas veces puede terminar siendo obsoleta, principalmente si se trata de productos con vencimiento cercanos o aquellos productos donde se trabaje por temporada. Si bien parte de los gastos incurridos para producir y mantener estos productos ya fueron mencionados en los puntos anteriores, algunas veces los productos obsoletos luego acarrearán adicionalmente gastos de desecho o derrame.

- Reducciones de margen. Con el objetivo de evitar la obsolescencia, es posible que se tomen decisiones de accionarlos con campañas promocionales que impliquen una reducción del margen
- Costos de “pérdida” de ventas. Corresponde a la ganancia no capturada de los productos que fueron solicitados por el comprador pero no se vendieron por falta de stocks, muchas veces también llamado Stock Out.
- Traslados a destinos equivocados. (Aronsson & Jonsson, 2016) “Costos extra de logística e inventarios producidos por traslados debido a envíos hacia destinos equivocados”
- Costo de oportunidad del capital en movimiento. El capital de trabajo permanece inaccesible perdiendo la oportunidad de utilizarlo para otros gastos/inversiones.

Otros autores presentan adicionalmente a los conceptos mencionados anteriormente, otros que no tienen una implicación tan directa y son, muchas veces, más complicados de dimensionar.

- Utilización ineficiente de los recursos de la empresa (Aronsson & Jonsson, 2016). Muchas de las decisiones de las empresas afectan la inversión y disponibilidad de capital, planificación de recursos humanos, etc. Están basadas en las proyecciones de ventas, tanto a corto como largo plazo. Tanto los errores como las variabilidades conocidas del forecast implican una dificultad mayor para tomar decisiones acertadas.
- Estrategias de promoción y marketing (Aronsson & Jonsson, 2016). En algunos negocios, promociones múltiples lanzadas en conjunto pueden resultar en la canibalización de ambos productos promocionados y terminar sin promocionar ningún sku. Integrar promociones y forecasts a nivel de distribuidor permite mejorar el flujo de bienes. Además, consigue obtener mejores resultados en términos de disponibilidad de stocks y tasas de llenado. De manera similar, mejoras en la habilidad de proyectar el impacto de los cambios de precios permitirá tener mejores ingresos y margen.

- Satisfacción del cliente. (Aronsson & Jonsson, 2016) Las pérdidas de ventas mencionadas anteriormente pueden afectar también la percepción del cliente sobre el servicio prestado por la empresa, con posibles implicancias negativas sobre futuros pedidos del mismo.
- Negociación de mejores términos con los proveedores. (Aronsson & Jonsson, 2016) Visibilidad sobre una proyección más acertada sobre las necesidades de recursos traerá beneficios también para la planificación de los propios proveedores, quienes tendrán un gran beneficio. Este beneficio puede ser captado por ambas partes a través de una mejor negociación de términos.

II.II. Variables que dificultan la preparación del forecast

Existen diferentes causas que generan errores en la proyección de demanda y estos muchas veces coexisten. Según Annastiina (Kerkkänen, 2008) existen 3 riesgos principales:

- Game Playing, se refiere al hecho que las personas que utilizan el forecast lo hacen para satisfacer sus propios intereses y muchas veces pueden sesgar la proyección
- Baja motivación, muchas veces el personal que prepara el forecast al no beneficiarse de la exactitud del mismo puede no verle ninguna ventaja y en consecuencia no alcanzar la motivación suficiente para buscar continuamente una mejora en el mismo
- Falta de habilidad, por último, también muy habitual en las empresas, el personal que prepara el forecast no cuenta con las herramientas o habilidades necesarias para lograr un resultado confiable

Los primeros 2 puntos pueden ser abordados mediante un análisis de los procesos y responsables que terminan preparando y ejecutando el forecast. Es recomendable revisar el proceso completo del S&OP.

El proceso de S&OP consta del refinado de los procesos de producción y demanda, en donde se repasa la base del forecast de demanda. Se busca consenso de los planes de producción y demanda y requieren de la responsabilidad y compromiso de cada una de las áreas involucradas para

asegurar que cada área ejecute tan de cerca como posible en lo que está siendo incorporado en los planes. Buscando de esta manera mejorar el accuracy del forecast, todas las áreas involucradas deben participar en la evaluación y revisión del desarrollo del forecast tanto como sea necesario (Lapide, 2006).

El tercer punto es donde se hace mayor foco en esa tesis, tratando de entender que modelos estadísticos y herramientas existen y son más adecuadas dependiendo de los recursos y tipo de demanda a analizar.

Sin embargo, existen varias limitaciones, tener las herramientas adecuadas significa también tener disponibilidad de datos históricos adecuados, lo cual, en muchos casos, es imposible corregir o recuperar.

Algunos de los puntos más comunes que no permiten contar con el histórico de la demanda correcta son:

- Falta de datos sobre stock Out.
- Falta de registro sobre promociones o pushes específicos nuestros o de la competencia.
- Falta de registro de eventos externos que afectaron las ventas en su momento.

Adicionalmente, en algunos casos, teniendo datos históricos confiables, existirán dificultades externas que harán afectar la predicción a futuro por dificultad de predicción de algunas de las variables que afectan la demanda.

Dificultad de predecir con precisión las variables que afectan la demanda

- Promociones nuestras y promociones de la competencia
- Temperatura, etc.(variables simples)
- Variables económicas que afectan el consumo de los hogares, en particular en argentina
- Existe una gran variedad que agentes externos que pueden tener un gran impacto en el contexto e influir de gran manera

sobre la demanda, el ejemplo más relevante de los últimos tiempos fue la pandemia (Covid 2019) que tuvo un impacto enorme en las proyecciones de la demanda en casi cualquier industria durante el 2020

Si bien algunos tipos de datos son más comunes y serán más fáciles de obtener, como las ventas netas, o históricos de algunas variables ajenas a la empresa que puedan afectar las ventas, como puede ser la temperatura o la cotización del dólar, existen varios puntos que suelen no ser relevados o proyectados debidamente y pueden dificultar la realización de un correcto forecast.

Por último, un factor importante a tener en cuenta es el efecto látigo (Bullwhip) de los cambios de demanda. Cuanto más intermediarios existan entre el cliente final y el punto donde se efectúa la medida de ventas que se utilizará para preparar el forecast, este efecto cobrará mayor importancia e incrementará cada uno de los impactos mencionados anteriormente. Cambios en la demanda no anticipados en la proyección de esta, pueden tener impactos muy negativos debido a este efecto.

La infraestructura de la cadena de suministro y la toma de decisión racional de quienes crean los pedidos en conjunto con las siguientes 4 fuerzas principales generan el efecto látigo: la actualización del forecast de demanda, las órdenes en lotes, la fluctuación de precios y por último el racionamiento y escasez de producto (Lee, Padmanabhan, & Whang, 1997).

MARCO EMPÍRICO

I. Introducción

Con el objetivo de mantener la confidencialidad de los datos presentados a continuación, se decidió cambiarle el nombre a la empresa que analizaremos, a la cual llamaremos Beers.

Beers es una empresa multinacional dedicada a la producción y venta de bebidas alcohólicas y no alcohólicas. Si bien compete en varias categorías, su principal negocio es el de las cervezas, siendo uno de los líderes en esta industria.

La empresa cuenta con presencia en Argentina donde cuenta con 8 plantas y más de 15 centros de distribución, más de 10 marcas de cervezas y una muy amplia red logística para llegar a cada punto del país con sus productos.

Si bien se trata de una empresa multinacional, Beers es una empresa que todavía maneja varios procesos descentralizados e incluso se tiene distintos niveles de tecnologización en diferentes países.

Durante este capítulo se aplicarán los conceptos analizados en el marco teórico para analizar el proceso de preparación del forecast e impacto del mismo.

II. Problema

Los principales problemas que motivaron a la realización de esta tesis corresponden a la baja tecnologización de los procesos de preparación del forecast de demanda de la empresa Beers, los bajos accuracy demostrados históricamente de este forecast, así como también las demoras en la entrega del mismo que mencionaremos en el siguiente capítulo.

Estos problemas resultan en costos relacionados con el sobre stock y la falta de stock principalmente, así como también los asociados con la alta variabilidad del accuracy como mencionado en capítulos anteriores.

En los siguientes capítulos se repasan los principales procesos que impactan la proyección del forecast para entender el impacto del accuracy del mismo y luego dimensionar los más importantes.

Luego se ahondará en las soluciones de forecast que podrían aplicarse con sus ventajas y desventajas asociadas.

III. Proceso de Forecasting

III.I. Planificación del forecast desde el área de Revenue Management

III.I.I. Preparación del forecast de ventas

Dentro de la dirección de Revenue management, el equipo de demanda realiza la preparación del forecast de ventas. Dentro de este equipo hay un analista designado tiempo entero a la preparación del forecast de cervezas y al seguimiento de las ventas del mismo negocio.

Este proceso se realiza enteramente en Excel, y comienza exportando los históricos desde los distintos sistemas de la compañía al cierre del mes.

Para obtener el forecast de ventas, una vez obtenidos los históricos a mes cerrado, se realiza el forecast estadístico sobre las ventas totales utilizando 3 inputs principales: el histórico de los últimos 3 meses (la tendencia), el histórico del mismo mes del año anterior y el valor de ventas definido en presupuesto. El analista define, en base a un criterio definido en el momento, cuánto peso asignarle a cada una de estas 3 variables. No se tiene una definición concreta del criterio a tomar para realizar esta ponderación, sino que se realiza arbitrariamente teniendo en cuenta principalmente de la evolución de cada una de estas variables, por ejemplo, si el patrón de la evolución de este año es similar al del último año o al definido en presupuesto.

Una vez obtenido el forecast estadístico de manera agrupada para la totalidad de los productos se realiza el desglose por marcas, calibres, regiones y canales. Para dicho desglose se considera la evolución del mix de productos de los últimos meses, es decir de qué manera se distribuyeron estos productos porcentualmente en los últimos meses para cada categoría mencionada.

En paralelo con este desglose, se enriquece el forecast considerando las acciones comerciales planificadas para el próximo mes, estas acciones pueden

tener efecto sobre marcas específicas, sobre regiones, canales o también a nivel de sku.

Al finalizar este enriquecimiento del forecast, se tiene el forecast de ventas que se envía al equipo de Logística y que será utilizado para planificar la producción y transporte de productos hacia las distintas zonas del país.

Tal como se mencionó previamente, para poder realizar el forecast el equipo de demanda descarga el histórico de ventas de los sistemas de la compañía al cerrar el mes, para luego procesar la información y entregar el mismo entre el cuarto y quinto día hábil del mes corriente.

III.I.II. Datos utilizados para el modelado

Cabe destacar que en este proceso de proyección de la demanda no se tiene en cuenta ninguna variable externa de la cual se tenga ya cierto conocimiento de su futura evolución. Solo se toma en cuenta datos históricos de las ventas y el presupuesto proyectado, el cual, a su vez es calculado en base a volúmenes de años anteriores, objetivos o estrategia de la empresa y rentabilidades de cada producto.

Por otro lado, el histórico de ventas utilizado para este modelo contiene los datos reales de ventas a mes cerrado, lo cual, tal como comentado en el marco teórico, muchas veces se tiene información sucia o errónea, y que no siempre resulta la más adecuada para la proyección estadística de la demanda. Solo a modo de ejemplo, no se tiene reflejado las pérdidas de ventas, acciones de ventas específicas o eventos externos que pueden haber ensuciado los datos.

III.I.III. Dificultades que enfrentan

El forecast final entregado tiene una granularidad de sku, canal y subregión con una visibilidad de 12 meses a futuro, pero con gran foco en el mes corriente.

Estas combinaciones de sku, canal y región suman un total aproximado de 7500 combinaciones, conformado por 145 sku de cervezas, 4 canales y 22 regiones (no todos los sku se venden en todas las regiones o canales). Esta cantidad de combinaciones, aun considerando el bajo nivel de complejidad del modelado,

hace que la proyección resulte muy pesada para ser preparada por una sola persona utilizando solamente Excel.

Ambos puntos mencionados en este capítulo terminan reflejándose en la precisión del forecast presentado. El equipo de demanda realiza un análisis sobre la evolución del accuracy a mes cerrado, se pudo observar que el mismo varía entre 70% y 75% en los últimos 3 años, habiendo entregado el forecast el 5to día hábil del mes corriente. Realizando benchmark sobre el accuracy del forecast obtenido dentro de la misma empresa, pero en otros países, se pudo observar que aquellos países que implementaron sistemas específicos para el armado del forecast tienen un accuracy que varía entre 80 y 85% aun entregando el mismo con una anticipación de 15 días a comenzar el mes.

Cuando hablamos de accuracy, nos referimos al MAPE (Error Porcentual Absoluto Medio), si bien es una métrica no muy recomendada para evaluar la precisión del forecast porque no toma en cuenta el impacto financiero de cada error, en Beers la principal métrica de seguimiento para evaluar esta precisión.

III.II. Planificación de la producción

III.II.I. Proceso de planificación de la producción

Una vez que el equipo de Revenue Management tiene el forecast enriquecido definido, este se comparte con el equipo de Demand Planning dentro de la dirección de Logística por correo. Con este forecast el equipo de Demand Planning genera el plan de producción por centro tomando en cuenta las restricciones específicas que tiene cada planta para luego compartir con distintas áreas de Supply Chain por correo.

Por un lado, el equipo de compras utiliza el plan de producción high level a 12 meses para armar el plan de necesidades a 12 meses permitiendo planificar la demanda de los insumos más críticos con anticipación, en particular latas, botellas y cajones.

Por otro lado, el equipo de planning de la central consultando información de los distintos sistemas de la compañía los stocks en plantas, stock en centros de distribución y producto en proceso, prepara un plan de producción por centro

detallado con un horizonte a 4 semanas y uno a mayor nivel a 8 semanas. El plan a 4 semanas para planificar los cambios de línea de cada planta y el de 8 semanas para permitir la planificación de insumos con anterioridad. Luego la central realiza seguimiento de los stocks, planes de producción y avance de ventas y se consolidan en Excel para dar visibilidad sobre la disponibilidad de productos. Este mismo plan de producción es utilizado por el equipo de Transporte para planificar los viajes desde las plantas hacia los centros de distribución, directamente hacia terceros (distribuidores o supermercados) o para exportación.

III.II.II. Problemas del proceso actual

Uno de los principales problemas que se observa en esta etapa del proceso se refiere al día del mes en el que el forecast estadístico y enriquecido es entregado al equipo de logística, esto sucede el 5to día hábil del mes corriente. Dado que la cerveza requiere cerca de 15 días para su cocción, el equipo de Planning utiliza el forecast del mes anterior para comenzar dicha producción, sin tener en cuenta posibles cambios o desvíos respecto de la estrategia originalmente planteada.

Otro de los problemas que se observan en este punto es la gran variabilidad existente entre lo planificado y las ventas reales generando grandes desvíos a medida que avanza el mes. Para mitigar mayor impacto el equipo de planning realiza ajustes en el plan de producción a lo largo del mes.

IV. Impacto de errores en el Forecast a lo largo de la cadena de suministro

En el marco teórico se mencionan diferentes tipos de impactos que puede llegar a sufrir una empresa cuando se tienen errores significativos en la predicción de las ventas. Muchos de estos aparecen al momento de analizar el accuracy del forecast de Beers y los costos que estos errores acarrearán.

El objetivo de este capítulo es repasar cómo afecta este bajo accuracy del forecast en cada eslabón de la cadena de suministro y dimensionarlos siempre que sea posible.

IV.I. Costos de reprogramación innecesarios

En el proceso actual el equipo de planeamiento de la producción se ve forzado a realizar cambios de programación de manera constante.

Como primer punto, el hecho de contar con el forecast el quinto día hábil, teniendo en cuenta que la producción del mes en curso para algunos productos comienza el día 15 del mes anterior, hace que sea necesario realizar un primer ajuste al recibir esta nueva información donde se ven actualizados nuevas proyecciones de ventas, así como también se reflejan posibles campañas que se realizarán en el mes corriente.

Luego las grandes diferencias entre proyección de ventas y la real hace que sea necesario realizar ajustes durante el mes en curso. Estos ajustes son muchas veces recomendados en cualquier industria, pero el bajo accuracy de las proyecciones hace que se realicen más ajustes de los deseados.

Dentro del impacto por cambios de línea es necesario considerar que cada cambio de formato o calibre requiere frenar la línea de producción entre 5 y 6 horas en promedio. A lo que debemos sumar los costos de la merma por purga de línea.

IV.II. Costos de producción

La producción de productos que no se venderán representa un gasto directo y en muchos casos resulta fácil calcularlos al asociarlos directamente con el producto no vendido.

Se tiene que el costo promedio de cerveza Main y cerveza Premium es de 1100 ars/HL y 4700 ars/HL respectivamente y que las cervezas premium representan un 23%.

La cantidad de producto que termina siendo obsoleto y va directo a derrame fue de cerca de 161700 hectolitros para el año 2020. Con lo cual se tiene un costo total de 188.493.690 ars.

IV.III. Reducción de margen

Este impacto resulta difícil de dimensionar ya que no se cuenta con información histórica donde se detalle cuando ciertas promociones fueron forzadas por tener exceso de producto y adicionalmente tampoco conocemos que impacto positivo se obtuvo de estas promociones. Sin embargo, sabemos que este tipo de promociones se realizan frecuentemente para evitar la obsolescencia.

IV.IV. Costos por pérdidas de ventas

Los costos asociados a las pérdidas de ventas son, muchas veces, uno de los impactos más difíciles de medir dado que pocas veces se registra adecuadamente esta información. En general cuando no se cuenta con cierto producto directamente se intenta reemplazar con otro o directamente se pierde la venta sin reflejar en ningún sistema. Para poder medir estos impactos es importante comenzar a registrar esta información

Sin embargo, podemos tomar como aproximación el siguiente cálculo utilizando el Service Level, este es un indicador utilizado para entender cuántos de los pedidos solicitados no pudieron ser entregados. Para el año 2020 este indicador fue de 2,03%, mientras que las ventas totales fueron 22.177.152 HL. Lo que significa que no se entregaron 459.708 HL y el motivo que más empuja este indicador suele ser la falta de stock en el centro de distribución al momento de la entrega.

Por otro lado, tenemos que la contribución marginal de nuestros productos es de cerca de 3400 ars/HL para productos main y 9800 ars/HL para productos premium representando 23% este último.

Esto resulta en una pérdida de más de 2 millones de pesos argentinos durante el 2020 por stock-out.

IV.V. Traslados a destinos equivocados

Dada la amplia red de plantas y centros de distribución con la que cuenta Beers a lo largo del país, al realizar las proyecciones de ventas se hacen estimaciones por zona, lo cual tiene impacto al definir la planta desde la cual se produce cierto producto, así como también el centro de distribución al cual se envía el mismo. Los errores en la proyección de ciertas zonas hacen que se necesite realizar envíos hacia destinos donde este producto sea requerido al cambiar la demanda.

IV.VI. Costo de oportunidad de capital en movimiento

El costo de oportunidad del capital en movimiento representa también un concepto importante dentro del caso de estudio. Por un lado, se tiene el costo de oportunidad de los productos que terminan siendo obsoleto y luego se tiene también que consistentes errores en el forecast fuerzan a una definición de políticas de stock más conservadoras obligando a la inmovilización de mayor cantidad de stocks.

IV.VII. Costo por derrame

Al momento de desechar el producto vencido, la empresa incurre en gastos de derrame. Cada hectolitro desechado tiene un costo de 18 usd. Durante el 2020 se derramó un total de 161700 hectolitros sumando un total de 2.846.000 usd en costos únicamente de derrame.

IV.VIII. Otros

Tal como se mencionó en el apartado teórico, existen otros impactos que son en la mayoría de los casos muy difíciles de dimensionar. Sabemos que estos impactos también afectan a la empresa Beers, pero no será posible realizar una estimación.

Entre estos se encuentran la utilización ineficiente de los recursos de la empresa, estrategias de promoción y marketing, bajas en la satisfacción del cliente, la posibilidad de negociaciones de mejores términos con proveedores.

V. Posibles soluciones de modelos predictivos

Tal como se vio al inicio del apartado empírico, la proyección de ventas es calculada utilizando principalmente la herramienta Excel como en muchas

empresas hoy en día. Esto limita mucho la capacidad de modelado de las series de tiempo y hace muy difícil incorporar todas las variables que sabemos que pueden afectar las ventas de manera directa.

Gran parte del modelado se basa en imitar la evolución vista el último año en el año actual utilizando el histórico de los meses o años anteriores como base para esta evolución y enriquecerlos con información sobre la estrategia de la compañía. Por lo cual se suele dejar fuera de este ejercicio mucha información histórica, así como también previsiones ya conocidas de variables que afectan las ventas de nuestros productos.

La baja tecnologización de dicho proceso no solo limita la capacidad de modelado, sino que además toma mucho tiempo al equipo de demanda y retrasa la entrega del forecast. Esta demora es clave a la hora de planificar la operación, teniendo en cuenta que la producción de cerveza tiene en muchos casos una media de 15 días de preparación.

Por esta razón es importante analizar diferentes soluciones para el armado de la misma.

El armado del forecast de la demanda requiere proyectar la evolución de las ventas para todos los productos de la compañía, aunque los SKU más críticos son la cerveza ya que se tiene un margen de acción mucho más reducido dados los tiempos de preparación del mismo. En total, dentro de la categoría cervezas se tiene un total de 145 SKUs. A su vez, la proyección se divide en subregiones y canales para tener claro en qué planta se producirá y hacia qué centro de distribución se tendrá que enviar. Luego de esta división se tiene un aproximado de 8500 combinaciones con 22 regiones y 4 canales de ventas (no todos los productos se venden en todas las regiones o canales).

Luego, es importante entender cómo afectan diferentes variables externas a la proyección de ventas. Variables de las cuales se puede obtener proyecciones independientes y que se puede llegar a encontrar cierta correlación con las ventas. Algunas de las variables conocidas que pueden llegar a tener alta relación con las ventas de cerveza son: temperatura, días festivos, humedad,

PBI, promociones propias o promociones de la competencia, eventos populares (mundial de fútbol por ej.), etc. Incluso una combinación de estas variables podría llegar a generar aún mayor cantidad de ventas, como ser temperaturas altas durante un día festivo.

De esta manera, considerando solamente las variables con relación más evidentes separando por región y canal, y relacionándolo para cada sku se obtienen proyecciones que resultan imposibles de realizar desde una planilla de Excel sin tener que realizar aproximaciones generales que simplifiquen los cálculos. Lo cual hace evidente la necesidad de herramientas más sofisticadas con mayor poder de cómputo que permitan obtener correlaciones de alta complejidad para cada combinación de sku y región.

V.I. Herramientas de Machine Learning Cloud

Una de las soluciones presentadas previamente considera la implementación de herramientas de planificación del forecast que utilizan Aprendizaje Automático para la preparación del modelo de forecast como es el caso de Amazon Forecast, que ofrece una herramienta que permite obtener correlaciones entre diferentes variables y generar la proyección con un poder de cómputo muy alto y con costos reducidos.

Amazon ofrece el servicio de Forecasting, al igual que muchos otros proveedores de servicios de forecast utilizando ML, con costos de manera variable, lo que permite comenzar a utilizarlo y validar su precisión sin entrar en costos altísimos. En este caso de ejemplo los precios de esta herramienta son:

Concepto	Precio
Forecasts generados	\$0.60 per 1,000 forecasts
Almacenaje de la información	\$0.088 per GB
Horas de entrenamiento	\$0.24 per hour

Tabla 1: Precios Amazon Forecasting. Fuente: <https://aws.amazon.com/es/forecast/pricing/>

Para nuestro caso de estudio, se tiene un total aproximado de 8500 combinaciones entre sku, canal y región, es decir 8500 series de tiempo. Adicionalmente sería útil tener proyecciones con diferentes percentiles, al menos una a 0.5 para obtener el forecast estándar, y uno mayor para tener un margen de seguridad adicional. Esto resulta en un total de 17000 series de tiempo.

Como capacidad de almacenamiento se puede estimar un total de 5gb lo cual es más que suficiente para almacenar una buena cantidad de datos históricos en formato .csv

Y finalmente se puede estimar un total de 20 horas de entrenamiento, aunque puede variar según la aplicación.

	Tarifa (usd)	Cantidad	Costo (usd/mes)
Forecasts	0.6	17000	10.2
Almacenamiento (GB)	0.088	5	0.44
Horas de entrenamiento	0.24	20	4.8

Tabla 2: Costo estimado Amazon Forecasting. Fuente: Creación propia

La suma de los gastos mensuales del servicio de Amazon Forecast resulta entonces en 15.44 usd al mes. A lo cual deberá sumarse el costo del equipo que monitoree la herramienta. Como se puede observar, el costo de este tipo de servicios es muy bajo. Vale la pena destacar que el caso de Amazon Forecast no es una excepción y existen otra variedad de servicios a similares costos, como el servicio de Google, Azure, etc.

V.II. Herramientas de forecasting On Premise

A modo de ejemplo tomaremos la herramienta Forecast Pro de Business Forecast Systems para el pronóstico de la demanda. Este proveedor ofrece licencias one-shot por cada usuario más un adicional del 15% del precio original para mantener la versión del mismo actualizada y con soporte técnico.

Precio	Usuarios
\$4,995	1
\$3,995/User	2-5

Tabla 3: Precio ForecastPro. Fuente: <https://www.forecastpro.com/>

Además, también ofrece capacitaciones completas para comenzar a utilizar la herramienta por un costo de 1800 usd online o 4500 usd por capacitaciones presenciales.

V.III. Consultoría especializada

Otra opción que podría analizarse es la de utilizar herramientas propietarias con consultoría especializada. Este tipo de servicios suele ser más caro que el resto de las herramientas mencionadas pero dada la magnitud de la empresa y los altos impactos que se obtienen de pequeñas diferencias de accuracy en el forecast (ya sea mejoras en el MAPE o el wMAPE), parece razonable al menos entender qué nivel de mejora se podría obtener para comparar con alguna de las herramientas de Machine Learning mencionadas previamente. Estos servicios cuentan con herramientas especializadas acompañados de soporte de especialistas con experiencia en el modelado de datos, poniendo mucho foco en la relación entre distintas variables y también en la preparación de los datos históricos para un mejor modelado. Un ejemplo de los proveedores de estos servicios puede ser el de Continente Siete, una empresa local con mucha experiencia específica en modelado de datos.

Una empresa de consultoría, que por temas de confidencialidad no nombraremos, ha cotizado para realizar este servicio, proponiendo una etapa de evaluación y prueba de concepto, por un total de 140.000 usd y un posterior costo mensual de 10.000 usd por el mantenimiento de la base de datos y modelado de la proyección de la demanda.

V.IV. Pros y Contras de cada solución

Lo primero que surge al analizar diferentes soluciones es la gran diferencia en los costos de estas alternativas, sin embargo, antes de apresurar una decisión sobre con qué alternativa debería avanzarse, se debe tener en cuenta las fortalezas y debilidades de cada solución.

Por un lado, las alternativas de Machine Learning como servicio ofrecen costos muy bajos, gran facilidad de uso, un alto poder de cómputo y alta tecnología para el modelado de proyecciones. Sin embargo, cabe destacar que para una buena proyección es muy importante tener muy buenos datos históricos, y esto, tal como se mencionó en el marco teórico, no siempre es fácil de obtener. La correcta preparación de estos datos es tan importante como el modelado de la proyección misma.

Por otro lado, la consultoría específica para el modelado de datos suele ser mucho más costosa, pero suele abarcar todos los aspectos importantes necesarios para obtener una buena proyección de datos, incluyendo la preparación de datos históricos, modelado de la proyección, pero también pueden acompañar con revisión de procesos que ayuden a una mejora continua. Además, el acompañamiento de personal especializado no es un punto menor al trabajar en esta ciencia que hoy es un ámbito desconocido para la empresa en cuestión.

CONCLUSIONES

Con este trabajo se ha intentado analizar el proceso de preparación del forecast, las herramientas utilizadas, el accuracy obtenido y el posible impacto asociado a este, para luego entender qué soluciones existen en el mercado y cuáles son las más recomendables para este caso.

Como primer punto se pudo ver que el proceso de preparación del forecast es muy artesanal y que no se utilizan herramientas con alto potencial o específicas para el modelado del forecast. Se pudo observar que, durante la preparación del forecast, se tiene los siguientes puntos de dolor:

- No se cuenta con una base de datos históricos limpia sobre la demanda real que permita un buen modelado de datos estadístico
- No se toman en cuenta la correlación entre las ventas y variables externas con proyección conocida que permitan una mayor precisión sobre la evolución de la demanda
- No se cuenta con un software específico que permita realizar modelados de alta complejidad

Al igual que la mayoría de las empresas actuales, se trabaja mucho con Excel, lo cual hace imposible realizar un modelado de alta complejidad para cada sku, canal y generar forecasts a partir de predictores relevantes.

Por otro lado, se han podido analizar algunos de los costos asociados a errores en la proyección de la demanda, donde se pudo observar que, dada la magnitud de la empresa, con altos montos de facturación, resulta que pequeños porcentajes de mejora en estos costos podrían llegar a tener un impacto en valores absolutos enormes, lo cual suma a la importancia de la utilización de buenas herramientas para el modelado de la demanda futura.

Por lo cual es altamente recomendable el estudio a detalle de la forma de preparación del forecast, justificando incluso la incorporación de sistemas complejos y con una estructura de personas especialistas o consultores para que lo soporten.

Soluciones Cloud de Machine Learning como las comentadas en este trabajo resultan una muy buena opción para un primer enfoque que permita acceder a algoritmos complejos para alcanzar un buen modelado a partir de datos históricos y de las correlaciones con todas las variables que puedan afectar este forecast, siendo herramientas de fácil acceso tanto a nivel económico como de conocimiento sobre estas herramientas.

Luego también sería recomendable entender qué nivel de accuracy pueden ofrecer otros softwares específicos para este trabajo en este caso puntualmente, siendo la opción de acompañamiento con consultoría específica la más recomendada. Esto permitiría enfocarse en el tratamiento de los datos históricos además de los puntos mencionados.

Bibliography

- Aronsson, H., & Jonsson, R. (2016). *Sales Forecasting Management*.
Gothemburg.
- Athanasopoulos, H. a. (2018). *Forecasting: Principles and Practice*. Melbourne.
- Basson, L. M., Kilbourn, P. J., & Walters, J. (2018). *Forecast accuracy in demand planning: A fast-moving consumer goods case study*. AOSIS.
- Céspedes, A. (2017). *Construcción de modelo de forecast para estimación de demanda en una empresa multinacional de retail*. UNIVERSIDAD TECNICA FEDERICO SANTA MARIA.
- Chollet, F. (2017). *Deep Learning with Phyton*.
- Glass, S. (2018). *AI and the Evolution of Demand Forecasting*.
- Harald Bauer, P. B. (2017). *Smartening up with Artificial Intelligence (AI)*. Digital McKinsey.
- Kerkkänen, A. (2008). *Demand forecasting errors in industrial context*.
Lappeenrant: Elsevier.
- Lapide, L. (2006). *TOP-DOWN & BOTTOM-UP*.
- Lara-Benítez, P. (2021). *An Experimental Review on Deep Learning Architectures for Time Series Forecasting*. Sevilla.
- Lee, H. L., Padmanabhan, V., & Whang, S. (1997). *The Bullwhip Effect in supply Chains*. MITSloan.
- Makridakis, S., Wheelwright, S., & Hyndman, R. (1997). *Forecasting Methods and Aplications*. Wiley.
- Marsland, S. (215). *Machine Learning: An Algorithmic Perspective, Second Edition*. CRC Press.

Oskar Blomgreen, S. E. (2016). *Improving Demand Planning and Inventory Management*. Sweden.

Polanco, A. E. (s.f.). *Series de Tiempo: Conceptos Básicos de Análisis de Series de Tiempo en el Dominio del Tiempo*.

Sarle, W. S. (1994). *Neural Networks and Statistical Models*. SAS Institute Inc.

Vandeput, N. (2021). *Data Science for Supply Chain Forecasting*. Belgium.