



TRABAJO FINAL DE MASTER

MODELO PREDICTIVO

DE GESTIÓN DE FLOTAS:

HORMIGONES LATINOAMÉRICA C.A

Grupo 3:

Noemi Cobo Diaz

Félix Wong Plasencia

Henrique Texeira Oliveira

Raul Trujillo Tagliafico

# Índice

[**Resume**](#_heading=h.ngjglll0l522)**n**……………………………………………………………………….…………………………………………….. 4

[**Introducción**](#_heading=h.gjdgxs)……………………………………………………………………………………….………………………. 5

[**Marco Teórico**](#_heading=h.n3vpe8hd09pe)……………………………………………………………………………………………………………... 8

* 1. [Importancia del Hormigón……………………………………………………………………... 11](#bookmark=id.le0z3b79vb61)
  2. [Hormigones Latinoamérica C.A……………………………………………………………..…. 14](#bookmark=id.hde1hrcw6oyg)

[**Glosario**](#_heading=h.c3avyw73lbbx)………………………………………………………………………………………………………………………. 17

[**Objetivos**](#_heading=h.3wen0hd6epuy)…………………………………………………………………………………………….…………….…….... 17

* 1. [Objetivo General…………………………………………………………………………](#bookmark=id.xdnr2zsbu4q0).………. 17
  2. [Objetivo Específico……..…………………………………………………………….…………….16](#bookmark=id.v6zfghv213m9)

[**Marco metodológico**](#_heading=h.rxya6akilsvt)……………………………………………..…………………………………………………… 18

* 1. [Sistema de Seguimiento de Flota…………………..………………………………………… 18](#bookmark=id.cg3xrzwguqu4)
  2. [Procedimientos y Conocimiento de Negocio……………..………………………………. 20](#bookmark=id.s57k3jg38u7)
  3. [Programa Diario……………………………………………………………..……………………… 21](#bookmark=id.1yu28h9ewvf6)
  4. [Implementación de un Modelo de Predicción………………………………..………….. 22](#bookmark=id.ap1wxjvu5u2c)
  5. [Limpieza y Tratamiento de los Datos……………………………………………………..… 23](#bookmark=id.3xma1nx4s5zv)
  6. [Cronograma de Actividades…………………………………………………………………….. 24](#bookmark=id.wyd07ioic39k)

[**Resultados**………………………………………………………………………............................................ 25](#bookmark=id.m1lp32xnj44e)

* 1. [Limpieza de Datos………………………………………………………………………………… 25](#bookmark=id.ix59g38psfia)
  2. [Análisis de Datos……………………………………………………………](#bookmark=id.230lwhisoql7)…………………...…. 26

[**Discusión de Resultados**..………...……………………………………………………………………………... 33](#bookmark=id.q74pq2ha3t7z)

* 1. [Conclusiones de Gestion…………………………………………](#bookmark=id.jml1nfb2fyuy)…………………………..…. 35

[**Reflexiones Finales**…………………………………………………………………………………….…………….. 37](#bookmark=id.k8iaz55ou763)

[**Referencias Bibliográficas**………………………………………………………………………………………... 4](#bookmark=id.majpc1rp8eu7)2

[**Anexos**](#_heading=h.84goy9f93u9q)……………………………………………………………………………………………………………………. 41

# Índice de Figuras

[Figura 01 - Experiencia de Compra de Concreto…………………………..………………...…………](#bookmark=id.41tg6rgs670k)..6

[Figura 02 - Tiempo de Ciclo………………………………………………..………………………….………](#bookmark=id.nax75l6de3lw)..15

[Figura 03 - Arquitectura de Solución………………………………………………………………...……...](#bookmark=id.9jvq2mith65p)20

[Figura 04 - Data Flow………………………………………………………..………………………….………..](#bookmark=id.qnd9deqzntba).21

[Figura 05 - Tiempo de Ciclo………………………………………..…………………………………….…….](#bookmark=id.1bnpjeoou1oz).23

[Figura 06 - Tabla Descripción de Variables………………………………….………………..……………](#bookmark=id.qsa3c0jk1pao)24

[Figura 07 - Cronograma de Actividades Gantt Chart……………………………..…….……………](#bookmark=id.m5k0v3yz6rv1)...26

[Figura 08 - Proporción de Data Apta para Análisis Estadístico……………..……………….……...26](#bookmark=id.fdueu4lwnyj3)

[Figura 09 - Retorno de la Inversión…………………………………………………...…………………](#bookmark=id.7irvvsfmz44m)…..38

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# Resumen

Hormigones Latinoamérica C.A es una empresa dedicada a la fabricación y distribución de hormigón premezclado, con sedes en 4 países de Latinoamérica. Actualmente posee una facturación anual de 1.2 millones de euros.

En aras de mejorar la experiencia del cliente, la organización realizó en 2018 una encuesta de satisfacción, señalando el retraso en los despachos como el factor que más influía en la percepción de calidad de servicio, costando anualmente a la empresa un 8% de facturación total en penalizaciones, siendo que aproximadamente un 10% de los despachos llegan fuera de tiempo estipulado.

El equipo de investigación conformado por Noemí Cobos, Felix Wong, Henrique Texeira y Raul Trujillo se planteó la meta de reducir el impacto de dichas penalizaciones en un 20% lo cual a su vez se traduce en un aumento del revenue anual de un 4%. Para ello se ha dispuesto el diseño y test de modelos de Machine Learning que permitieran optimizar los tiempos de descarga del hormigón.

El estudio se basó en un dataset obtenido de reportes de la operación durante un período de dos años. Dada la naturaleza de los datos, los modelos se basaron en árboles de decisión en todas las configuraciones posibles, logrando reducir el tiempo de descarga de producto en un 50%.

Dicha información será incorporada al proceso de programación para reducir la ventana de tiempo de entrega y por ende aumentar la eficiencia operacional de la organización y a su vez la satisfacción de los clientes.

Palabras Claves.

Supply Chain, Programación Logística, Machine Learning, Algoritmos, Arboles de Decisión, Satisfacción del cliente, Ultima Milla, Transporte

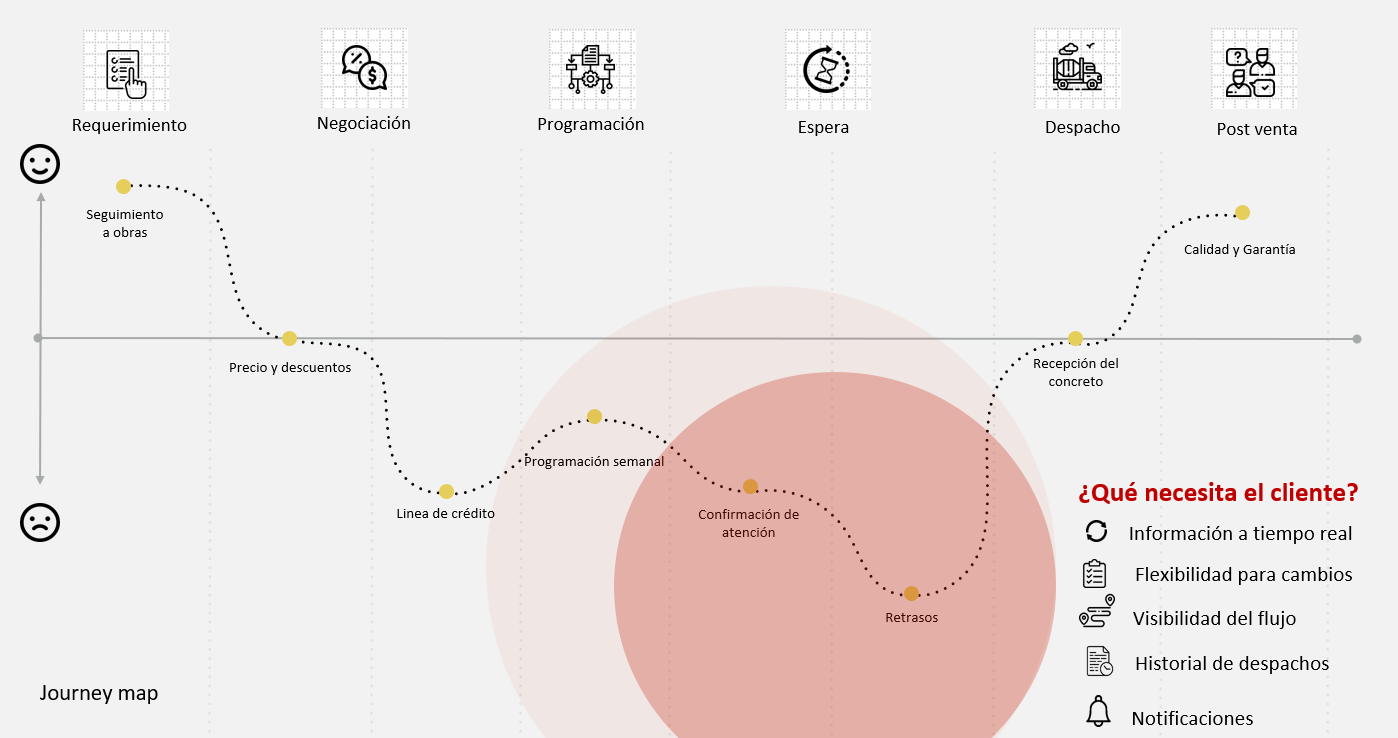
# Introducción

La administración de rutas de despacho ha sido desde el principio de la revolución industrial uno de los aspectos más relegados en lo que a procesos de optimización se refiere. Hoy en día con un mundo cada vez más globalizado, y con el movimiento de mercaderías desde cualquier parte del planeta, en industrias que son sensibles a mermas o penalidades por fallas en entrega, cobra vital importancia hacer uso de las tecnologías a nuestra disposición para afrontar los escenarios cada vez más complejos a los que se enfrenta la última etapa del proceso de supply chain. Dichas industrias, por las características intrínsecas del producto que comercializan, implican tener un control exhaustivo de los tiempos de entrega para garantizar que el despacho del producto o servicio se está haciendo en condiciones óptimas de calidad. Por ello, la incorporación de modelos de predicción de tiempo de despacho permite obtener un gran volumen de datos para procesar y analizar mediante la aplicación de tecnologías como la Inteligencia Artificial o el Machine Learning. De esta forma, las industrias son capaces de segmentar y granular las características de sus clientes para posteriormente clusterizar datos y obtener insights que hacen posible definir el perfil de los clientes para tomar decisiones estratégicas ya sea para ahorrar costes o para mejorar procesos productivos identificando áreas de mejora e implementando políticas de eficiencia.

Tales beneficios, pueden llegar a representar la diferencia entre un modelo exitoso de negocio o una pérdida sustancial del beneficio. Algunos ejemplos de estas industrias son: farmacéuticas, artículos perecederos como lácteos, víveres, hortalizas y frutas frescas, repuestos de la industria aeronáutica y naval, bienes médicos, trasplante de órganos y cualquier otra aplicación médica de urgencia, suministro de alimentos y provisiones en caso de desastres naturales o cualquier otra industria cuyo producto o servicio que comercializa está sujeto a unos tiempos de entrega específicos. Todas estas industrias enfrentan desafíos complejos a diario en el manejo de las variables internas y externas y es indispensable tener la operativa del servicio controlada y gestionada adecuadamente ya que una mala praxis puede generar retrasos en la entrega, estimación errónea de costes, disconformidad del cliente con el servicio y, en supuestos específicos, la pérdida de vidas humanas.

El presente proyecto de investigación está enfocado en el análisis de *supply chain* de una compañía de suministro de hormigón en Latinoamérica. La decisión de poner el foco del trabajo de investigación en esta industria y en dicho territorio se debe a la gran disponibilidad de datos con los que se cuentan para el análisis de las diferentes variables que influyen en los tiempos de despacho (se dispone de una fuente de datos diarios de monitoreo del proceso de despacho de hormigón de 2018 a 2020).  Esta tesis que se desarrollará en los próximos apartados, surge a raíz de las conclusiones de una encuesta de satisfacción de servicio que la compañía realizó a 50 clientes de los que dispone en cartera. Dicha encuesta se realizó en el año 2017 (noviembre) y tenía como objetivo medir el grado de satisfacción de los clientes en cada etapa de su experiencia de compra.

Figura 1 - Experiencia de compra de concreto



De las respuestas obtenidas, se determina como una de las conclusiones principales que los clientes no están satisfechos con el tiempo de despacho de los pedidos que realizan a la empresa. Esto es, que el tiempo de despacho fue uno de los puntos críticos al obtener una puntuación muy baja y por ende una de las causas que genera mayor insatisfacción al cliente. Dicha inconformidad obedece a varios motivos, pero fundamentalmente cada retraso implica al cliente extra costes en términos de personal no ocupado, el cual es dispuesto para recibir la carga y que en la mayoría de las oportunidades se trata de personal externo a la obra contratado para dicho único fin.

Para mejorar el indicador, se estudiarán las diferentes fases operativas del despacho y se enfocará en la mejora de la eficiencia de aquella fase donde el tiempo de ejecución dependa en mayor medida de variables conocidas o controlables por la empresa. En el ámbito de esta investigación se ha determinado que son 5 las fases o tiempos de los que se compone el proceso, siendo el “tiempo de descarga” donde estará centrada la investigación.

El objetivo de este estudio tiene dos vertientes: Por un lado, el objetivo técnico que sería el desarrollo de un modelo de predicción para poder conocer con un día de antelación el tiempo de descarga del camión destinado a una obra. Este dato proveerá información valiosa para la asignación de flota disponible al día siguiente y poder organizar los recursos materiales y humanos de forma más eficiente.

Por otro lado, el objetivo de negocio tiene como último fin y, a consecuencia del objetivo anterior, mejorar la satisfacción del cliente y, a su vez, generar ahorros importantes a la compañía tras mejorar la operativa de despacho haciéndola más productiva.

El fin último de esta investigación es usar el modelo de predicción de tiempo definido como herramienta de soporte para la planificación, control y despacho del hormigón. Los beneficios esperados son incrementar el nivel de servicio de cara a los clientes internos y externos, mejorar el control y seguimiento de despacho, y reducir el porcentaje de despachos impuntuales, lo cual inevitablemente se traduce en mayor revenue para la organización

Por todo lo expuesto anteriormente, el estudio se enfoca precisamente en el área de programación de despachos, la cual es la pieza fundamental de todo el proceso.

# 

# Marco Teórico

En el mundo actual, prácticamente todas las empresas que venden productos y materias primas tienen un componente de despacho en su proceso de comercialización. Existe a día de hoy un ambiente cada vez más globalizado en donde las entregas en tiempo y forma son fundamentales por lo que cada acción e imprevisto tiene un impacto en los resultados. En la medida que tenemos acceso a mayor cantidad de datos de múltiples fuentes, la percepción personal de los involucrados y su intuición o experiencia, si bien valiosa, son aspectos que cada vez toman menos relevancia, dando paso a considerar solo lo que se puede medir y analizar. Esto permite a las empresas anticiparse al futuro, lo que se traduce en mantener un control más efectivo sobre costes inesperados, aumentar la productividad y en última instancia la rentabilidad de la organización.

La cadena de suministro genera a diario una cantidad inmensa de datos tanto estructurados como no estructurados que sólo pueden explotarse gracias a la inteligencia artificial. La logística se basa en redes físicas y digitales que no pueden ser optimizadas por humanos debido a su alta complejidad. Por ello, el objetivo de la inteligencia artificial es transformar los comportamientos reactivos en proactivos, los manuales en automáticos y los estandarizados en personalizados (Gesing, Peterson, & Michelsen, 2018).

Según Elizabeth Rodrigues (Management & Innovation (Núm. 24) · TIC · abril 2020, “La digitalización es un motor de cambio que está transformando la sociedad y que afecta a todos, desde consumidores a empresas y trabajadores. Y todos los cambios vienen asociados a nuevas oportunidades, pero también a riesgos difícilmente controlables”.

En el ámbito de la gestión de flotas, existen muchos campos en donde la optimización de procesos puede ejercer una influencia directa en el resultado. Algunos de ellos son los mantenimientos preventivos programados, la optimización de rutas y tiempos de despacho, la monitorización de indicadores en tiempo real a través de la telemetría e IOT e incluso el factor humano de la ecuación.

Marcelo Leporati y Manuel F Morales así lo corroboran en su disertación Inteligencia Artificial en la gestión de cadenas de suministro”, al afirmar que ¨La capacidad de vincular e integrar las distintas áreas y actividades de la cadena de suministro en un sistema que permita predecir fallos y defectos utilizando la IA tiene un gran valor para todo el proceso¨.

La Inteligencia Artificial (IA) y el Machine Learning (ML) son cada día más factores determinantes para evitar situaciones costosas e indeseadas. Así mismo la predicción basada en data y el monitoreo realtime pueden generar modelos de predicción que permitan proyectar diferentes escenarios en función de las decisiones a tomar. Las compañías que ya la utilizan adquieren una ventaja competitiva de enorme relevancia respecto aquellas que no las aplican.

La Inteligencia Artificial e Internet de las cosas (IoT) permiten a las empresas ahorrar grandes cantidades de dinero en diversas áreas. No solo se trata de que un software basado en modelos de predicción de datos pueda avisar al programador de la flota sobre las consecuencias que implicaría el incumplimiento de un mantenimiento programado. Hoy es perfectamente posible entrenar modelos de IA para que recomienden el mejor momento para que una flota pase por mantenimiento en el taller, con independencia de las indicaciones generales de los fabricantes

A pesar de que el concepto de mantenimiento predictivo no supone una novedad, el desarrollo tanto de las tecnologías de obtención y almacenamiento de datos como el de las aplicaciones de Machine Learning, han contribuido a crear una nueva perspectiva sobre este término. De este modo, la contribución de datos procedentes de muchas fuentes es tratada con algoritmos más complejos que han permitido una reducción en los costes de mantenimiento de entre un 10-15 % (Breunig, Kelly, Mathis, & Wee, 2016).

En los últimos años se está alcanzando la capacidad computacional sobre los datos necesaria, en cuanto a almacenamiento y en cuanto a procesamiento, para integrar todos los datos contextuales, incluso de terceros, en un solo modelo de Inteligencia Artificial. Esto se consigue gracias a software con una arquitectura flexible y Cloud, con la que se procesan enormes volúmenes de datos críticos en tiempo real.

Los modelos de datos con los que se establecen predicciones de IA se configuran y mejoran aprendiendo por sí mismos de la definición de varios escenarios posibles. Para ello se tiene en cuenta tanto la data proveniente de los componentes del vehículo como las condiciones de utilización de la unidad. Es decir, que se puede distinguir entre, por ejemplo, el diferente uso que se realiza de un vehículo cuando se emplea para recorrer largas distancias, los recorridos intermedios o los recorridos de escasos kilómetros, pero con abundantes paradas.

Aquí se logran progresos fundamentales en materia de gestión y productividad de las flotas de vehículos, ya que es posible alcanzar el dimensionamiento ideal de ésta. ¿Cómo responder a picos altos de demandas de servicios de transporte? La IA indica cuál debe ser el tamaño de la flota en cada momento, durante todo el año y durante hitos específicos y permite tomar medidas como, por ejemplo, aumentar el número de vehículos con modelos financieros flexibles en esas fechas destacadas.

Por otra parte, la IA es una vía muy útil para mejorar la rentabilidad de las flotas. Es decir, que permite ofrecer un mayor volumen de servicios sirviéndose de los mismos recursos disponibles en ese momento, o el mismo volumen de servicios con menos recursos.

Los softwares que utilizan IA y ML para la optimización de rutas basan su operativa en la recopilación constante de data y en su análisis, con una lectura también continua de cómo afecta esto al servicio de la flota y al rendimiento. Se trata de soluciones tecnológicas que procesan millones de datos y toman miles de decisiones al mismo tiempo que se ejecutan. Muchas de ellas, incluso, optimizan y recalculan las decisiones de enrutamiento en tiempo real y de acuerdo a nuevos inputs. Dicha tarea sería mucho más lenta, propensa a errores y costosa si tuviera que ser afrontada por parte de personal humano.

La mayor trazabilidad posible de cada transporte y de cada envío es una necesidad y un reto para todas las empresas del sector, incluidas las de la logística de producción. Conocer al detalle el viaje que realiza cada producto o cada vehículo, en el caso de flotas de servicios, es cada día más importante, porque está en juego que todos los eslabones de la cadena funcionen. También, en especial en la última milla, que tanto el operador como el cliente conozcan, en todo momento, dónde se encuentra el producto y por qué diferentes fases han pasado. En resumen, data y transparencia. (Sheffi, Y., Saenz, M. J., Rivera, L., & Gligor, D. (2019))

La IA y las aplicaciones del machine learning en la logística y el transporte son capaces de adelantar, con sus modelos, predicciones de cómo se desplazarán los envíos y los vehículos en los distintos tramos de la cadena, estableciendo previsiones antes de que el envío se realice. Por ende, es un valor agregado complementario a la habitual lectura que se hace en tiempo real de las trazas medibles y analizables de envío de una mercancía.

Anticiparse a las necesidades de los clientes es una gran ventaja competitiva dentro de la cadena de suministro. Con la IA es posible, en tiempo real, recopilar datos, guardarla e interpretarla. Con esos datos se alimenta un modelo que es capaz, con el correspondiente entrenamiento, de detectar patrones y realizar previsiones. Esto es muy útil para ajustar el servicio a los valles y a los picos de demanda, y para actuar de acuerdo con ellos.

Con el uso de Inteligencia Artificial y Aprendizaje Automático puede predecirse con alta precisión algo tan importante para la circulación de la flota como las condiciones del entorno en el que van a trabajar; desde las condiciones climatológicas hasta el estado de las vías y del tráfico.

La Inteligencia Artificial y el Machine Learning suponen una importante disrupción en la gestión de flotas. El paso de un modelo de administración basado en una mentalidad transaccional a otra fundamentada en aplicaciones con algoritmos cognitivos representa una gran oportunidad de mejora para las compañías que en su día a día operan con vehículos. Las mejoras en eficiencia y la identificación de ahorros se logran gracias a modelos que recopilan y almacenan datos de manera masiva, y que después proyectan distintos escenarios posibles. De esta forma se toman las mejores decisiones de gestión, se minimizan las incidencias negativas y los errores, y además se tiene una capacidad de reacción inmediata y óptima ante cualquier eventualidad relacionada con la operativa.

El desarrollo de la IA y el ML y su aplicación de la gestión de flotas está en gran auge, por eso los próximos años serán testigo del crecimiento de sus aplicaciones. Cada vez es más factible ejecutar análisis de múltiples fuentes de datos en tiempo real, así como capturar y analizar más datos. Con ello se establecen correlaciones directas entre las acciones de gestión y la eficiencia operativa. A medida que la tecnología de IA avanza, los gestores tienen una mejor visión y un mayor control de su flota, que puede ser monitorizada hasta el detalle y de detectar, de manera anticipada, incidencias. La información recopilada se puede procesar y analizar para conocer de manera automatizada cuál es el comportamiento normal de los vehículos y de los conductores, evitando acciones que resulten costosas e ineficientes e incluso situaciones de fraude.

La implementación de la IA no es un proyecto con un alcance claramente detallado y temporal y con recursos definidos, es más bien un camino que las compañías deben comenzar a andar, experimentar, adaptar, corregir e iterar a medida que el entorno lo vaya requiriendo (Leporati, Morales 2019).

Importancia del hormigón

Entre las diversas soluciones para enfrentar los retos de la construcción de infraestructura a nivel mundial destaca el desarrollo de materiales de alto desempeño, el uso de tecnologías innovadoras que ayuden a optimizar los procesos constructivos y el prolongar la vida útil de las estructuras, sin lugar a dudas el hormigón es la mejor alternativa y el que mayor demanda tiene a nivel mundial por su variabilidad de aplicaciones. Puede ser empleado para construir puentes, represas, edificios y otra gran variedad de obras civiles. La importancia del hormigón en los proyectos de infraestructura radica en su versatilidad, desarrollo de tecnologías que lo han llevado a límites insospechados en su desempeño, usos y aplicaciones. Actualmente, es el material de construcción más ampliamente utilizado en el mundo con una producción mundial cercana a los 13,000 millones de m3 por año (Ceballos 2019). Las ventajas que hacen del hormigón un material imprescindible para la construcción de los grandes proyectos y le permiten responder ágilmente a los principales desafíos de la infraestructura son:

1. Capacidad de resistir una gran variedad de condiciones de exposición extremas durante su vida útil, gracias a su alta durabilidad y resiliencia.
2. Es un material local y de alta disponibilidad que puede ser fabricado en cualquier parte del mundo, lo que ayuda a optimizar los costos y reducir la huella de carbono.
3. Sus propiedades estéticas permiten innovaciones arquitectónicas y flexibilidad en su diseño.
4. En aplicaciones de infraestructura (cimentaciones, túneles, etc.) el uso del hormigón es insustituible.

El proceso de fabricación del hormigón se inicia con la explotación de los yacimientos de materia prima (piedra caliza), en tajo abierto. El material resultante de la voladura es transportado en camiones para su trituración, los mismos que son cargados mediante palas a cargadores frontales de gran capacidad. La trituración de la roca, se realiza en dos etapas, inicialmente se procesa en una chancadora primaria, del tipo cono que puede reducirla de un tamaño máximo de 15 m hasta los 25" cm. El material se deposita en un parque de almacenamiento. Seguidamente, luego de verificar su composición química, pasa a la trituración secundaria, reduciéndose su tamaño a 2 mm aproximadamente. El material triturado se lleva a la planta propiamente dicha por cintas transportadoras, depositándose en un parque de materias primas. En algunos casos se efectúa un proceso de prehomogeneización. La siguiente etapa comprende la molienda, por molinos de bolas o por prensas de rodillos, que producen un material de gran finura. En este proceso se efectúa la selección de los materiales, de acuerdo al diseño de la mezcla previsto, para optimizar el material crudo que ingresará al horno, considerando el cemento de mejores características.

El material molido debe ser homogeneizado para garantizar la efectividad del proceso de clinkerización mediante una calidad constante. Este procedimiento se efectúa en silos de homogeneización. El material resultante constituido por un polvo de gran finura debe presentar una composición química constante. La harina cruda es introducida mediante sistema de- transporte neumático y debidamente dosificada a un intercambiador de calor por suspensión de gases de varias etapas, en la base del cual se instala un moderno sistema de precalcinación de la mezcla antes de la entrada al horno rotatorio donde se desarrollan las restantes reacciones físicas y químicas que dan lugar a la formación del clinker. (ASOCEM, 2019) El intercambio de calor se produce mediante transferencias térmicas por contacto íntimo entre la materia y los gases calientes que se obtienen del horno, a temperaturas de 950 a 1, 100ºC en un sistema de 4 a 6 ciclones en cascada, que se encuentran al interior de una torre de hormigón armado de varios pisos, con alturas superiores a los cien metros. (S. NAVARRO.)

Básicamente para la obtención del hormigón se tienen como materias primas el cemento, el agua y los agregados (grava y arena). Comúnmente también se emplean aditivos para modificar propiedades como su durabilidad, trabajabilidad y resistencia. De la misma manera, es posible adicionar fibras metálicas o sintéticas para aumentar su resistencia a la flexión, en aplicación como revestimientos de túneles con hormigón lanzado. (Gutiérrez, 2019)

El hormigón premezclado es una masa de cemento, agua y agregados (arena y grava) que comienza a fraguar inmediatamente al entrar en contacto con el agua creando una reacción química, que visto bajo el microscopio, da la apariencia de comenzar a salir raíces de las partículas de cemento, creando un entramado fuerte entre la grava y arena. Este proceso de “enraizado” se da durante las primeras 2 horas de mezclado, por lo que al ir girando dentro de la olla se va fracturando este entramado para evitar el endurecimiento antes de llegar a la obra. Es por ello que es de suma importancia planear y cronometrar los tiempos de recorrido y entrega del producto para evitar su baja resistencia.

De acuerdo a lo que se establece en la norma internacional ASTM, el material es mezclado cuidadosamente en la planta mediante rigurosos estándares de calidad que le da las propiedades requeridas para la aplicación específica del cliente, con la consistencia y revenimiento correcto para facilitar el vaciado y con la adecuada resistencia y durabilidad. Entre los beneficios adicionales del material se pueden mencionar su uniformidad en aspecto, color y resistencia, la aceleración del ritmo de trabajo y el bajo costo de mano de obra.

Las propiedades de esta mezcla plástica pueden ser adecuadas a las necesidades de cualquier aplicación que se desee y ambiente externo. Éste es un material de construcción muy adaptable que puede desempeñar su función por largo tiempo y con un mínimo mantenimiento, siempre y cuando sea utilizado correctamente para las especificaciones con que fue solicitada por el cliente.

Una planeación adecuada es la clave para una perfecta entrega en la obra. Para garantizar la entrega del producto en condiciones óptimas de calidad, debe de evitarse en todo momento los retrasos en el sitio. Para ello es fundamental contar con el personal necesario y con el equipo adecuado para que la logística marche en correcto orden. Generalmente los primeros 30 minutos después de efectuarse la mezcla no representa ningún riesgo de endurecimiento temprano siempre y cuando se haya mantenido en agitación dentro de la olla. Posteriormente se cuenta hasta con 4 horas después para vaciar, compactar y efectuar el acabado. No es necesario mencionar que pasado este tiempo el material pierde sus propiedades y debe ser desechado, con todo lo que ello implica en términos de costes y productividad.

Hormigones Latinoamérica C.A

Hormigones Latinoamérica C.A es una empresa privada que forma parte del sector industrial y minero, cuya finalidad es comercializar y producir cementos, cal, bloques, hormigón y agregados. Sus operaciones están enfocadas en Latam, y cuenta con ocho plantas ubicadas en Chile, Ecuador, Colombia y Perú. Cuentan con una flota aproximada de 122 mixers de hormigón y 21 bombas distribuidas entre las ocho sedes, dependiendo del diseño operativo mensual.

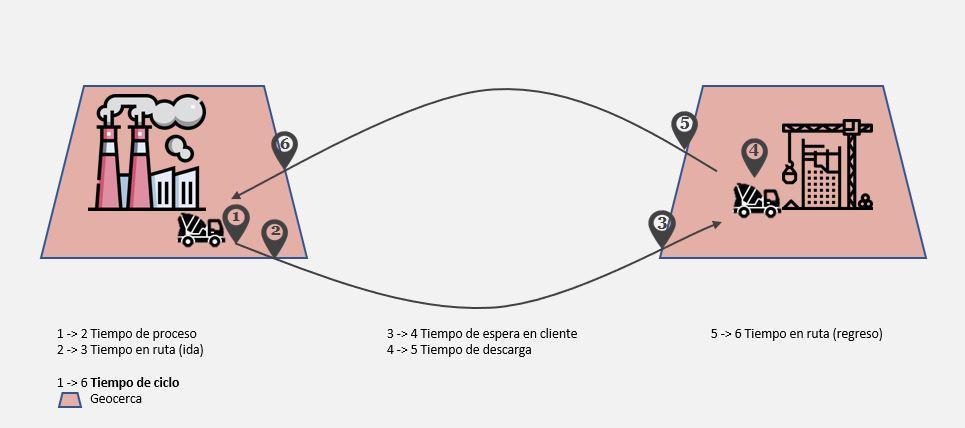
El Proceso comercial se efectúa a través de su empresa subsidiaria comercial de nombre Distribuidora Hormigones LATAM que posee una amplia red de distribuidores asociados en la región para ofrecer una mejor atención a sus clientes.

En el caso específico de la organización Hormigones LatinoaméricaC.A cuando una obra (cliente final) solicita concreto para iniciar operaciones, primero se negocian los términos de despacho en un contrato y posterior a ello, se procede a registrar la solicitud en un programa semanal. Tres días previos a la fecha de despacho solicitado se realiza una visita técnica para verificar el avance de obra y de todo estar conforme se procede a brindar la conformidad de atención. En este punto la solicitud ya es considerada un pedido y es registrado en el programa diario de despacho asignado para dicha fecha.

En la actualidad, toda la flota es monitoreada por un único Operador de GPS. Por ende, se ha desarrollado una integración entre el módulo de ventas (SD – SAP) y la plataforma del proveedor actual, en la cual replica automáticamente la información del programa de despacho y toda la data maestra necesaria: Cliente, Punto de expedición, cantidad a despachar, destinatario, geocerca (long, lat, R), placa, chofer, etc.

Durante todo el trayecto de la unidad, adicional al seguimiento en línea, se construye en línea un reporte “tiempo de ciclo” que muestra todas las fases del despacho. Como se observa en la imagen existen seis eventos generados según reglas de negocio, y la resta de los tiempos de ocurrencia de cada evento dará origen a cada fase.

Figura 2 - Tiempo de Ciclo



**Tiempo de proceso**: Inicia desde el abastecimiento de concreto al mixer, continua durante las pruebas de calidad, lavado de mixer y finaliza al salir de la geocerca. Este proceso es controlado por la empresa, tiene una media de 15 minutos y poca variabilidad al ser un proceso bastante estándar.

**Tiempo en ruta:** (Ida/Regreso): Es el tiempo que le toma al mixer desplazarse entre la planta y la obra, tanto de ida como vuelta. El tiempo está cronometrado desde que sale de la geocerca de planta hasta ingresar a la geocerca del cliente, y viceversa.

**Tiempo de espera**: Comprende el tiempo “muerto” o “de espera” para que el mixer sea atendido por el responsable de obra. Este tiempo depende netamente del cliente, y es una estadística importante frente a problemas de calidad y futuros reclamos. Este inicio desde el ingreso a la geocerca de obra y finaliza con la activación del sensor de descarga del mixer.

**Tiempo de descarga:** Es el tiempo que toma abastecer de concreto una determinada estructura en la obra. Y se contabiliza desde que se activa el sensor de descarga de concreto hasta que el mixer sale de la geocerca del cliente.

**Tiempo de ciclo:** Equivale a la suma de las fases, y nos indica cuánto tiempo le toma a un mixer realizar el despacho volver y estar nuevamente disponible para ser reasignado.

Es importante recalcar que los contratos de compra/venta de concreto incluyen cláusulas de penalización por incumplimientos, entre ellas, la impuntualidad con los tiempos de entrega. Si se toma en consideración la naturaleza del trabajo que se realiza esto tiene un sentido fundamental, pero también representa una clara amenaza a la rentabilidad de la empresa. En el momento de realización de este estudio, aproximadamente uno de cada 10 despachos llega con retraso y son penalizados, lo cual cuesta a la empresa un 1% del ingreso total por penalizaciones o multas.

Adicionalmente a esto, es posible que si el despacho no se lleva a cabo en la ventana de tiempo de 2.5 - 3.4 horas, el producto corra riesgos de calidad y deba ser desechado en su totalidad al haber perdido sus propiedades fundamentales. En el año, cerca del 1% de despachos realizados pasan por este escenario debido a una deficiencia en el proceso de crear el programa de producción.

# Glosario

**Cemento:** Es una mezcla de clinker con yeso, usado como materia prima para elaborar hormigón.

**Hormigón:** También llamado concreto, es la mezcla realizada a partir de cemento, agua, piedra, arena y aditivos plásticos, usada extensamente para la construcción.

**Mixer:** Vehículo pesado para y diseñado para el transporte de hormigón en estado fresco.

**Tiempo de ciclo:** Proceso completo de entrega de hormigón, comprendido desde el abastecimiento de concreto en un mixer, hasta que el mismo regresa a planta posterior a la entrega.

**Tiempo de descarga:** La cuarta fase del tiempo de ciclo, equivalente al tiempo empleado para descargar el concreto en obra y retirarse de la geocerca.

**Geocerca:** Una geocerca es una cerca virtual o perímetro alrededor de una ubicación física. Como una cerca real, una geocerca crea una separación entre esa ubicación y el área que la rodea

**Data wrangling:** Transformar los datos crudos en datos utilizables y que agregan valor

**Datos estructurados:** Archivos de tipo texto que se suelen mostrar en filas y columnas

**Datos no estructurados:** Conglomerado masivo y desorganizados de datos en diversos formatos.

**GPRS 3G:** Servicio General de paquetes vía radio, permite la mensajería instantánea, utilizado en mensajes SMS, correos electrónicos, entre otros

**Inteligencia Artificial:** Programa de computación creados para copiar o simular algunas operaciones consideradas únicamente humanas

**IOT:** Internet of Things es la captación de datos a través de sensores, softwares y otros, con el propósito de conectar e intercambiar datos entre dispositivos y sistemas a través de la internet

**Machine learning:** El estudio de algoritmos computacionales que mejorar automáticamente a través de experiencias y datos

**Programador:** Persona que programa la logística de los despachos

**Telemetría:** Sistema de medición de magnitudes físicas que se transmite a un observador lejano

**Trigger:** en español gatillador, es un evento que desencadena una acción, generalmente la ejecución de algún código de programación para iniciar o continuar un flujo.

**Variable dependiente:** Hace referencia a la variable cuyo valor está determinado por una o más variables independientes. Generalmente representada por la letra Y o la expresión f(X)

**Variable independiente:** Hace referencia a la variable cuyo valor no es afectado por otras variables. Generalmente representada por la letra X

**RMSE:** De las siglas en inglés “Root Mean Square Error”, en estadística, raíz del error cuadrático medio (en español), es una medida que compara la precisión entre valores reales y estimados.

**MAE:** De las siglas en inglés “Mean Absolute Error”, en estadística, Error Medio Absoluto, es una medida que compara la precisión entre valores reales y estimados. La diferencia entre RMSE y MAE es la sensibilidad frente a la data evaluada, RMSE >= MAE

# Objetivos

En concordancia con lo expresado en la parte introductoria del presente proyecto de investigación, el desencadenante de este proyecto de investigación fue la aplicación de una encuesta de satisfacción de servicio que la empresa hormigonera llevó a cabo con 50 de sus clientes finales. La principal conclusión de dicha encuesta es que el departamento de planificación es la pieza fundamental del proceso y la que más puede agregar valor al tener a su disposición una herramienta que le permita optimizar su función. Es en dicha área funcional donde se enfoca el objetivo principal de la investigación, así como los objetivos específicos que enumeramos a continuación

Objetivo General

Desarrollar un modelo de predicción de tiempo de descarga de hormigón a un día visto que provea de una herramienta confiable y fidedigna al departamento de planificación para reducir la incidencia de incumplimientos en tiempos de despacho de caras al cliente, con la finalidad de aumentar el nivel de satisfacción de los clientes y optimizar el revenue anual de la organización.

Objetivos Específicos

* Determinar las variables independientes más influyentes en la predicción del tiempo de descarga de producto.
* Determinar la existencia o no de otras variables independientes adicionales que permitan a futuro mejorar la efectividad del modelo.
* Estimar, con base al histórico de datos que se disponen de la totalidad de los clientes, el impacto generado por mejorar la predicción en el tiempo de descarga
* Demostrar mediante la aplicación del modelo a un dataset de prueba que la implementación permite en efecto mantener por debajo del 8% los casos de despacho impuntual del producto.
* Demostrar mediante un ejercicio práctico con la data histórica, la efectividad financiera de la aplicación del modelo y el ahorro en metálico que su implementación podría proveer a la empresa
* Presentar las posibles mejoras a futuro del sistema y las posibles aplicaciones que se podrían implementar, así como también los requerimientos técnicos de dichas soluciones

# Marco metodológico

Tal y como se expresa en el apartado de objetivos, la hipótesis de esta investigación es que a través de la implementación de un modelo de regresión se puede predecir con un porcentaje de certidumbre de +- 20% el tiempo de descarga (único factor que depende en un 100% de la empresa Hormigones Latinoamérica) y que este dato permitirá disminuir el gap a 8% de despachos con retraso y sus respectivas consecuencias económicas para el negocio.

El primer paso para responder esta pregunta de investigación radica en entender cuál es la situación actual y dinámica de operaciones de la empresa. ¿Cuál es el tiempo estimado actual de descarga usado? ¿y el error de predicción sobre el mismo?

En segundo lugar, es realizar un modelo de predicción del tiempo de descarga que incluya variables independientes adicionales al formato actual que permita brindar una mejor predicción.

En tercer lugar, podemos realizar una comparación estadística de medias para determinar si el modelo obtenido es significativamente mejor o no.

Finalmente, con los resultados del análisis anterior de ser positivos, estimar si con el modelo obtenido podemos cubrir los objetivos trazados.

La estrategia para responder esta pregunta de investigación está basada en 3 pilares: (1) un sistema integral de seguimiento de flota, que proporcione, almacene e interprete los datos de posicionamiento geográfico de cada mixer durante el despacho, (2) procedimientos y conocimiento de negocio, y (3) la implementación de un modelo de predicción que pueda aprender de la data (1) y reglas de negocio (2) para predecir el tiempo de descarga.

Sistema de seguimiento de flota

La implementación de una plataforma de seguimiento de flota tiene dos grandes frentes, la sensorización de los vehículos, que consisten en equipar a cada mixer con un GPS, un plan de datos 3G/4G y sensores de descarga que permitan generar los datos para monitorear los eventos deseados en el tiempo de ciclo, y el desarrollo de plataforma donde se volcará el conocimiento de negocio para definir reportes que generen valor.

Por lo general, los sensores de GPS brindan información de número de serie de equipo, matrícula de vehículo asignado, posición geográfica, hora, velocidad de desplazamiento, dirección y estados del dispositivo (encendido/apagado/error).

El GPS establece una comunicación satelital para determinar la posición determinada por latitud y longitud. Cabe resaltar que esta conexión difiere a la conectividad móvil celular brindada por las empresas de telecomunicaciones. En decir, la respuesta del satélite suele ser bastante alta y precisa.

Internamente el GPS incorpora un plan de datos móvil, lo que permite la transferencia de datos entre el dispositivo y las bases de datos de los operadores de GPS. La frecuencia de transmisión puede ajustarse dependiendo de la necesidad, solo debemos tener en cuenta que a mayor el flujo de información, más elevado será el consumo de datos y por ende más caro mantener el modelo. Para este caso, los GPS envían la información cada 60 segundos.

Para asegurar la correcta transmisión de datos, es relevante llevar indicadores de transmisión. Cada GPS debe enviar 1440 (1min \* 60min/hora \*24 horas/día) registros por día. De tener una disponibilidad de datos inferior al 97%, se debe brindar mantenimiento.

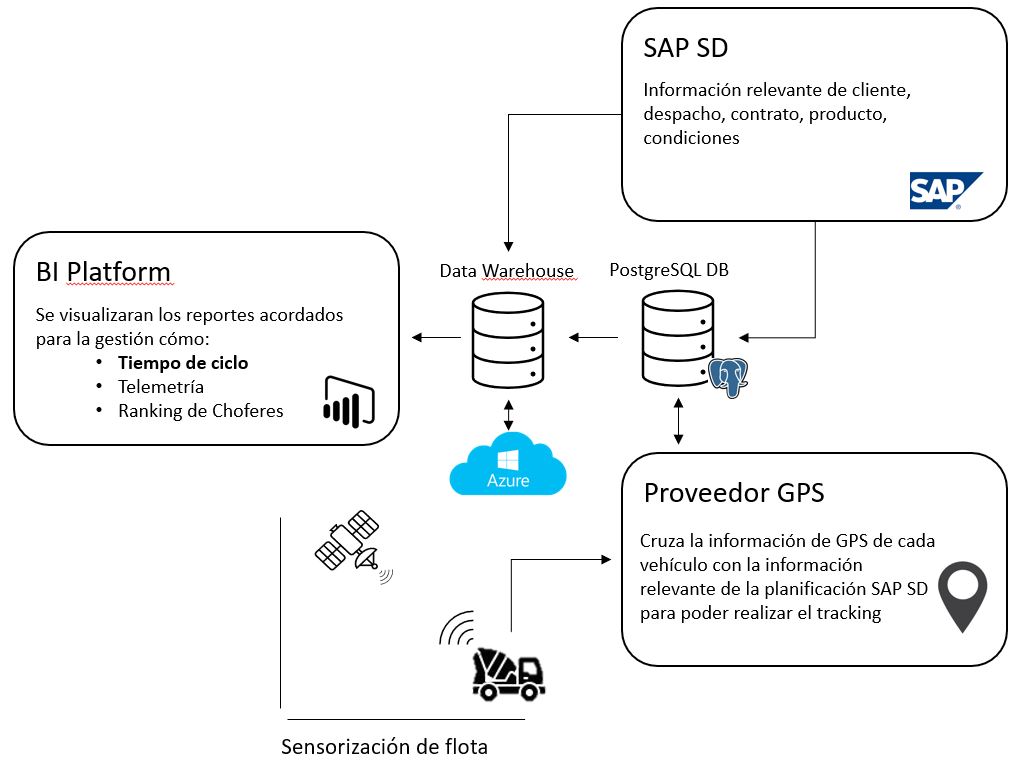
Dada la arquitectura de comunicación, se presentan problemas por falta de cobertura de red 3G/4G, es por ello que el GPS contiene una memoria interna, que, ante la falla de envío de datos, almacena los registros pendientes y cuando se encuentre en una zona con señal, enviará la información. Si bien se pierde el seguimiento en tiempo real, no se pierde la data para un análisis estadístico posterior.

La instalación de un equipo GPS no conlleva mayor complejidad, solo debe ubicarse en un lugar seguro donde no pueda ser manipulado y tener una conexión eléctrica. Adicionalmente, el equipo GPS suele ser escalable, en el sentido que puede agregar lecturas de otros sensores en la transmisión de datos.

Los sensores externos, en este caso el sensor del evento de descarga permite identificar el sentido de giro del trompo en el mixer. Un giro en sentido horario indica mezcla, y un sentido antihorario indica descarga.

Es importante tener en cuenta las condiciones y lógica de lectura de los sensores y equipos IoT, ya que ante fallas o errores de lectura podemos justificar y tomar acciones más adelante al interpretar la data. Por ejemplo, el GPS en movimiento con vibración en un mixer tiene un rango de precisión de +/- 20 metros, o en el caso puntual de sensores inductivos o magnéticos usados para determinar el sentido de giro, pueden enviar falsas alertas ante la presencia excesiva de partículas ferrosas que podemos encontrar en el polvillo de construcción durante una obra. Mitigar estos problemas abarca mucho entendimiento técnico, situacional y propio de la solución, por lo cual no se ahondará en este estudio en este sentido.

Figura 3- Arquitectura de Solución



El desarrollo de la plataforma de seguimiento comprende la integración del ERP SAP de la empresa a través de una base de datos externa (por seguridad) en PostgreSQL. Desde aquí el proveedor puede tomar la información de planificación con todos los datos necesarios (como cliente, latitud, longitud, estructura, nombre de obra, productos, cantidad, matrícula, número de despacho, etc.) y cruzarlos en su plataforma con los datos generados por la lectura de los sensores.

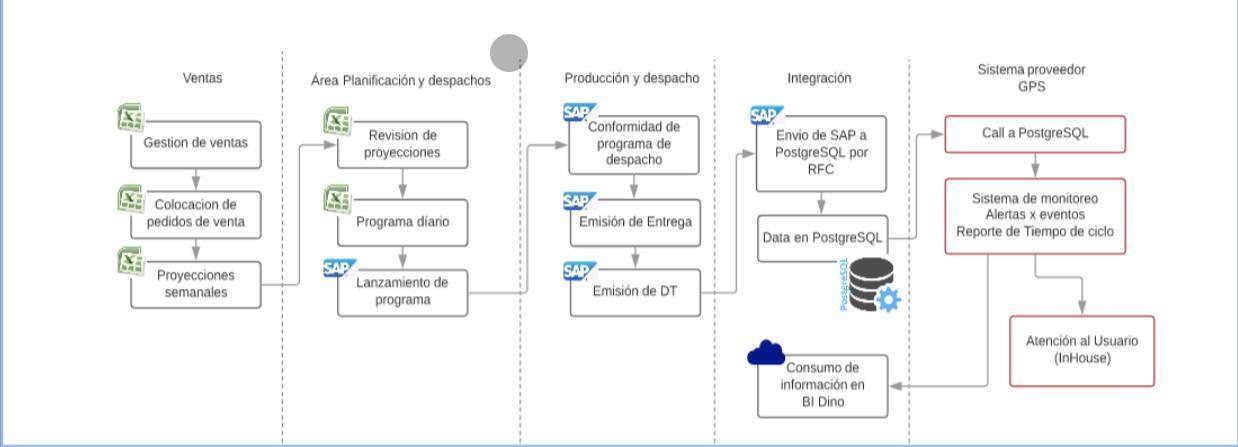
Es en esta plataforma propia del operador de GPS que tiene la lógica de interpretación de los datos GPS, capacidad computacional para procesar la información, interpretarla y presentarla. Muchas de las reglas de negocio son traducidas en líneas de código para poder obtener los reportes necesarios que generen valor, como el “Tiempo de ciclo” y el detalle de los 5 tiempos que lo componen, alertas de calidad por tiempos prolongados, estadística de impuntualidad de despachos, etc.

Procedimientos y conocimiento de Negocio

El procedimiento de venta y despacho de hormigón comienza cuando el representante comercial visita al cliente final para coordinar el volumen a despachar, tipo de hormigón a utilizar, las características físicas de la obra, el tipo de estructura, y la posición geográfica de la obra.

Toda esta información relevante es recopilada en un contrato en SAP e ingresada en un programa semanal de proyección de ventas.

Figura 4- Data Flow



Posteriormente un asesor técnico de ventas visitará la obra días previos al despacho para corroborar la información suministrada, validar el estatus de avance de la obra, y definir con más detalle los requerimientos técnicos de la descarga.

Una vez cargada la información en el sistema, todas las condiciones están dadas para que el programador, previa coordinación con el cliente, proceda a agendar todos los despachos de un día específico con una antelación de 24-48 horas (este proceso es llamado “Programa diario” y se detalla más adelante). Este proceso se realiza de manera telemática mediante una llamada y/o correo electrónico al cliente, con la finalidad de que los trabajadores estén prestos a recibir el producto. La falta de coordinación en este punto generalmente se ve reflejada en el “tiempo de espera” por motivos del cliente, el cual afecta directamente a la calidad del concreto si es muy prolongada.

Una vez lanzado el programa de producción y despachos, este debe ser aceptado por el jefe de planta, quien corrobora que se tenga todo lo necesario para llevarlo a cabo, caso contrario se realizan las coordinaciones para reprogramar los despachos.

Cuando estamos próximos a abastecer el hormigón dentro de los mixers, es el momento donde los documentos de despacho son generados, Documento de transporte (DT) y Guía de transporte (documento legal). La emisión de estos documentos son el trigger para iniciar el seguimiento de la unidad. La información del despacho junto con la matrícula del vehículo asignado es compartida con el proveedor de GPS, y posteriormente tratada en su plataforma donde cruzan los datos de sensores y lógica de negocio para obtener los reportes.

Dado que el comportamiento físico del hormigón varía según el tiempo, la temperatura, humedad relativa, los parámetros de calidad son bastante rigurosos durante el despacho. Estos pueden tener un impacto al interpretar los resultados, puede darse casos que los mixers no completen el recorrido y regresen a planta para realizar un reajuste de formulación, o realicen estos ajustes en obra incrementando algún tiempo determinado.

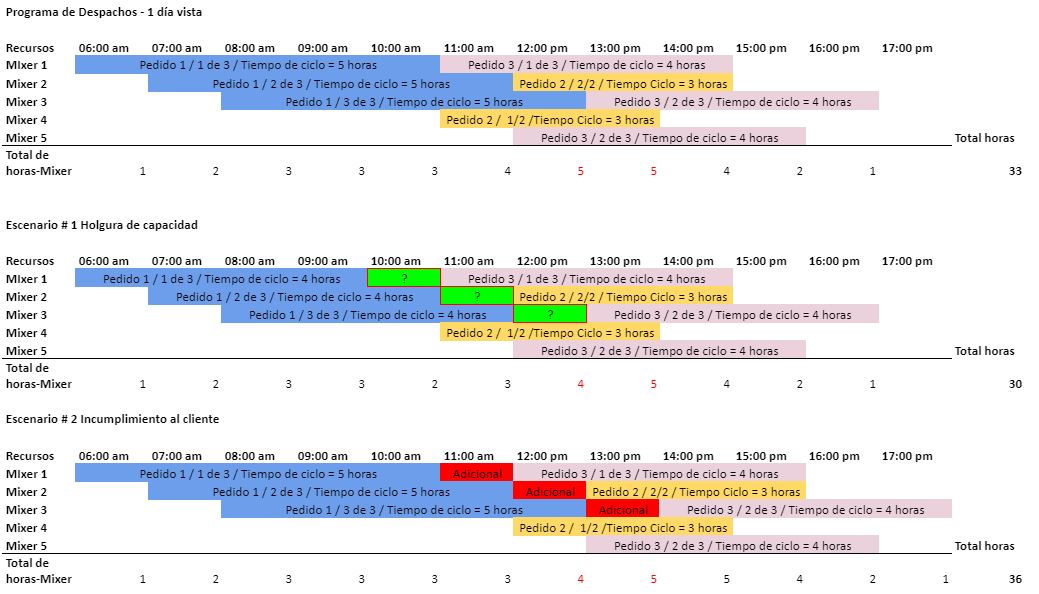
Es importante recalcar que los tiempos mencionados en el marco teórico son calculados de diferencia de hora de registro entre las coordenadas del hito A al hito B, hasta completar el ciclo, lo cual puede generar valores nulos o errados por motivos de mantenimiento de IoT, la dinámica de la operación, zonas sin cobertura celular, falsas descargas, mala respuesta del satélite, pérdidas de paquetes de información durante la integración, etc.

Se debe tener un entendimiento completo de la infraestructura de la solución de IoT y de la dinámica de negocio para poder identificar qué registros deben ser aceptados, cuales presentan errores, y que medidas se deben tomar para tener calidad y disponibilidad de los datos.

Programa Diario

Para detallar el impacto que presenta la optimización del tiempo de descarga en el programa de despacho y por ende en una acción que satisfará al cliente, se presentan los siguientes casos

En la imagen se aprecian un programa de despachos recomendado para el día siguiente. En él se detalla que se atenderán 3 pedidos (casillas de colores) a lo largo del día (eje X - Horas del día) con 5 mixers diferentes según su disponibilidad (eje Y - Recursos). El tiempo comprometido que tiene cada mixer está representado por casillas de un mismo color y obedecen al “tiempo de ciclo”.

Figura 5- Tiempo de Ciclo

Dada la predicción, el día de despacho se pueden tener 3 escenarios posibles:

* Que la predicción se cumpla, y el horario propuesto se cumpla a cabalidad.
* Que el tiempo estimado haya sido muy por encima del real, generando holgura de capacidad, y disminuyendo la productividad del negocio.
* Que el tiempo estimado haya sido muy por debajo del real, teniendo que dedicar más tiempo que el planificado, afectando al despacho programado siguiente, se puede comprar que el pedido 2 y 3 iniciaban despachos a las 11:00 am, sin embargo, se tuvo que posponer 1 hora dada la demora en la primera obra.

Implementación de un modelo de predicción

Para afrontar esta etapa se aplicarán herramientas de data wrangling para asegurar la consistencia de la data recabada durante el período relevante para el presente estudio, que abarca desde el 01 de agosto de 2019 hasta el 30 de diciembre de 2020, incluyendo todas las variables pertinentes posibles.

Posteriormente mediante un proceso de “features selection” seleccionaremos aquellas variables que agregan mayor cantidad de valor a la predicción, dejando de lado aquellas que no aportarían correlación estadística al modelo.

Luego probaremos con diversos modelos predictivos hasta obtener el mejor resultado de predicción posible. Y finalmente, compararemos estadísticamente el nuevo modelo de predicción vs el método actual.

Limpieza y tratamiento de los datos:

La naturaleza de los sistemas de recogida de información antes descritos permiten concluir que el número de datos nulos o errados existen, por lo que una estrategia de eliminación y normalización de los mismos es pertinente para contar con un dataset que permita alimentar los modelos de predicción deseados. Para efectos de facilitar la comprensión del lector se enumeran las etapas de la limpieza de los datos a continuación:

1. Identificación y cuantificación de los valores nulos: se propone contabilizar los valores nulos a través de la función isnull de pandas en todo el dataframe y declarar porcentualmente la estadística de los mismos
2. Eliminación de registros según reglas de negocio.
3. Definición de funciones de cálculo de latitud y longitud a distancia a través de la función Haversine
4. Conversión de variables categóricas en variables numéricas (dummies) para la posterior alimentación del modelo.
5. Visualización previa de la data a través de curvas estadísticas (Desviación Estándar, Varianza).
6. Representación visual de las posibles variables independientes y su correlación en pares.
7. Selección de variables con mayor importancia para el modelo de predicción.
8. Probar diferentes modelos de predicción hasta hallar el que presente mejor rmse.

En la siguiente tabla se establece la relación existente de las variables independientes a imputar al modelo.

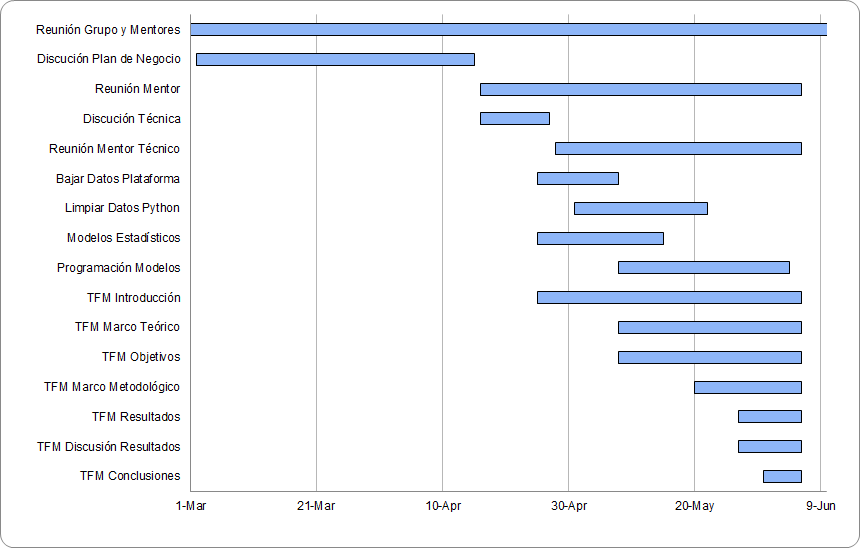
Figura 6- Descripción de Variables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Concepto | Nombre variable | Descripción | Tipo de variable | Impacto en el modelo |
| Distancia | WERKS | Centro de distribución | Cualitativa Nominal | Cada centro pertenece a un área geográfica distinta, donde el clima podría afectar |
| WERK\_LAT | Latitud Centro Distribución | Cuantitativa Continua | Usada para el cálculo de Distancia |
| WERK\_LON | Longitud Centro de Distribución | Cuantitativa Continua | Usada para el cálculo de Distancia |
| OBRA\_LAT | Latitud de Obra | Cuantitativa Continua | Usada para el cálculo de Distancia |
| OBRA\_LON | Longitud de Obra | Cuantitativa Continua | Usada para el cálculo de Distancia |
| DISTANCIA | Kilómetros entre Centro de distribución y Obra | Cuantitativa Continua | Correlacionado con el tiempo que estará en ruta la mezcla. |
| OBRATIPO | Clasificación de obra por tamaño según cantidad despachada | Cualitativa Nominal | Ante mayor cantidad de despachos históricos, menor error de predicción |
| Características del despacho | ID\_MODALIDAD\_2 | Despacho Directo / Bombeado | Cualitativa Nominal | El bombeo requiere un equipo adicional para poder trasladar el hormigón a zonas altas y por ende mayor tiempo de preparación |
| V\_ENTREGADO | Cantidad despachada | Cuantitativa Discreta | Directamente proporcional al tiempo empleado de descarga |
| ID\_ESTRUC | Código de estructura a ser vaciada con hormigón | Cualitativa Nominal | Cada estructura requiere de una técnica de abastecimiento distinta que involucra más tiempo. |
| Características del producto | ASENT | Asentamiento o Slung del hormigón | Cualitativa Nominal | Medida que refleja el grado de fluidez del hormigón |
| Tiempo | FINSEMANA | Si es sábado o domingo | Cualitativa Nominal | Se estima que los fines de semana se trabaja más lento, dado que hay menos mano de obra |
| FINMES | Último día del mes | Cualitativa Nominal | Se estima que los fines de mes hay menos mano de obra |
| FRANJA\_HORARIA | Identificador: madrugada, mañana, tarde, noche | Cualitativa Nominal | En los turnos diurnos se trabaja más rápido y eficiente que en los turnos de la tarde donde el personal presenta cansancio |

Cronograma de Actividades

Para desarrollar el siguiente proyecto hemos establecido el siguiente cronograma de actividades:

Figura 7- Cronograma de Actividades Gantt Chart



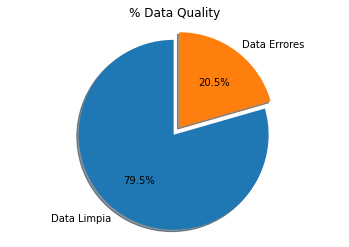
Resultados

La información obtenida a través de la plataforma de GPS, con la cual se ha realizado este estudio, está comprendida entre el 1 de agosto del 2019 al 31 de diciembre del 2020.

Limpieza de datos

En dicho periodo, se presentan 45.683 registros de despachos realizados de hormigón. De los cuales, solo el 79,5% es apto para ser usado en el estudio. El 20,5% es excluido del estudio ya que presentan datos nulos para la variable dependiente o tiempos irreales para el entendimiento del negocio.

Figura 8- Proporción de data apta para análisis estadístico



Reglas de negocio para depuración de data:

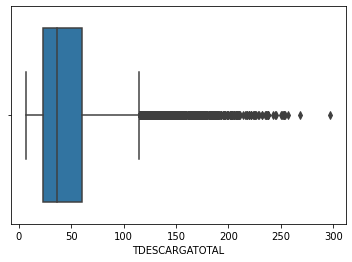
* Volumen de despachos por cada camión debe estar entre 0.5 y 8 m3.
* El tiempo de descarga debe estar en el rango entre 6-300 minutos.
* La máxima distancia para atender un despacho son 200 km.
* Excluir los escenarios de plantas dedicadas o plantas in-situ en obra.
* Un despacho no debe tener un tiempo de ciclo superior a los 400 minutos.

Concluida la limpieza de registros erróneos, se cuenta con 36.298 registros aptos para proceder con el análisis de datos.

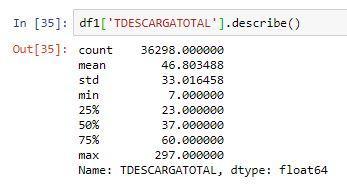
Análisis de datos

Comportamiento de variable dependiente: Tiempo de Descarga

Distribución del Tiempo de Descarga Diagrama de Cajas Tiempo de Descarga

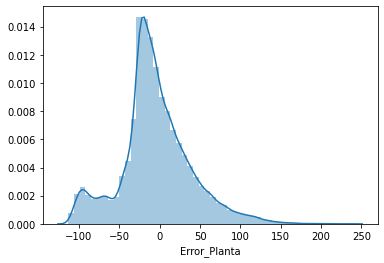
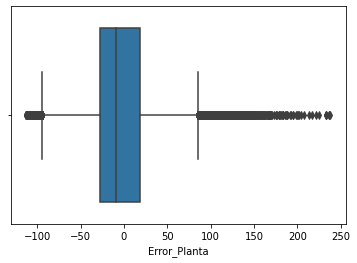


Resumen Estadístico - Tiempo de Descarga

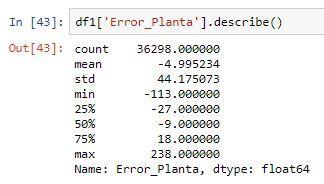


Comportamiento del error de la variable dependiente:

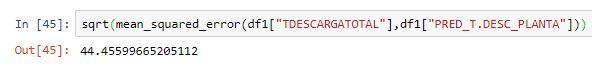
Real (GPS) - Pronóstico (método actual basado en T.prom según estructura)



Distribución del Tiempo de Descarga Diagrama de Cajas Tiempo de Descarga

Error de Predicción Método Actual Planta 

Raíz del error cuadrático medio actual (rmse): 44.45 min



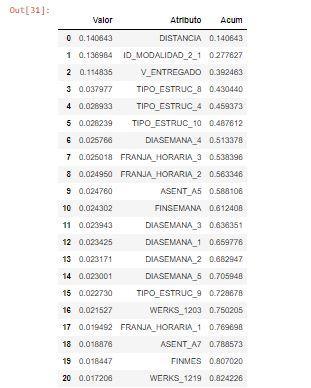
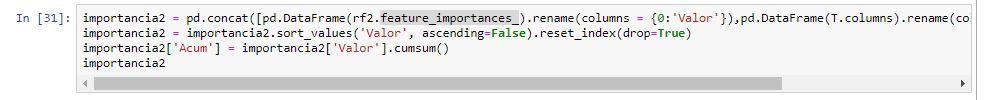
Después de presentar todas las variables independientes, se procede a revisar la correlación entre estas para evitar la colinealidad. De presentarse alguna correlación significativa (>0.6) entre las variables independientes, se debe eliminar aquella variable independiente que tenga menor correlación con la variable dependiente.

Dado el mapa de calor, podemos observar que ningún par de variables independientes presenta correlación alguna superior a 0.6

Paso siguiente, se procede a identificar la importancia de cada variable independiente para el modelo, también llamado “Feature Selection”. Para ello, creamos un modelo de tipo “random forest regressor” apta de entrenamiento y visualizamos el atributo importancia “feature\_importances”. Como se puede observar, se han ejecutado dos modelos, donde el modelo con mejor score es RF2 con un score de 0.73366

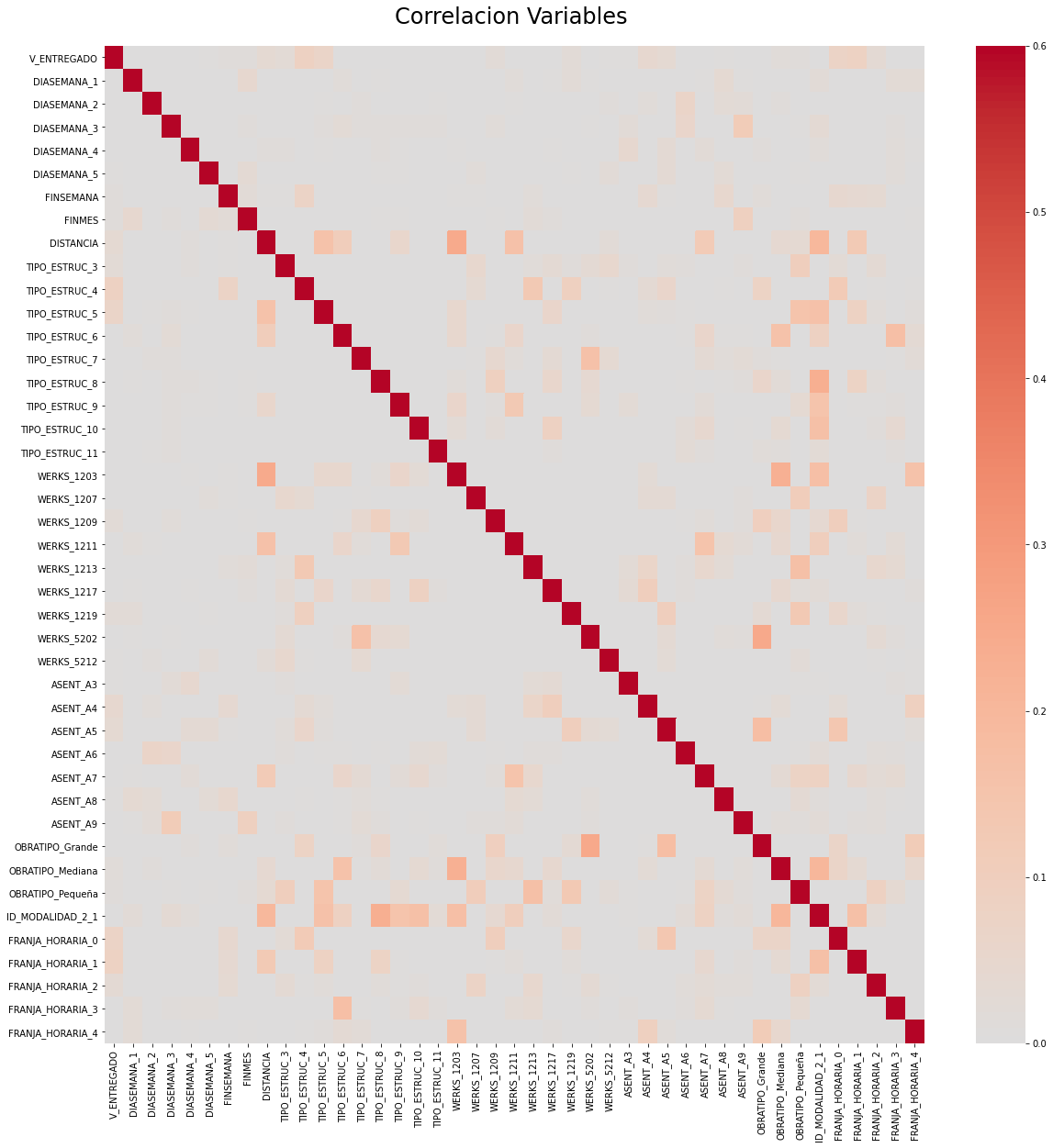


Se procede a llamar al atributo “feature\_importance” ordenado de mayor a menor y adicionando una columna acumulada para identificar las variables que impactan en el 95% de la predicción del modelo.



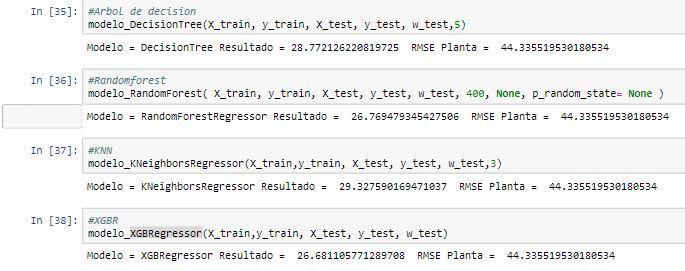


Finalmente, se procede a crear diversos modelos predictivos con las variables independientes que acumulan hasta un 95% de Feature\_Importance\_ del modelo RF2.

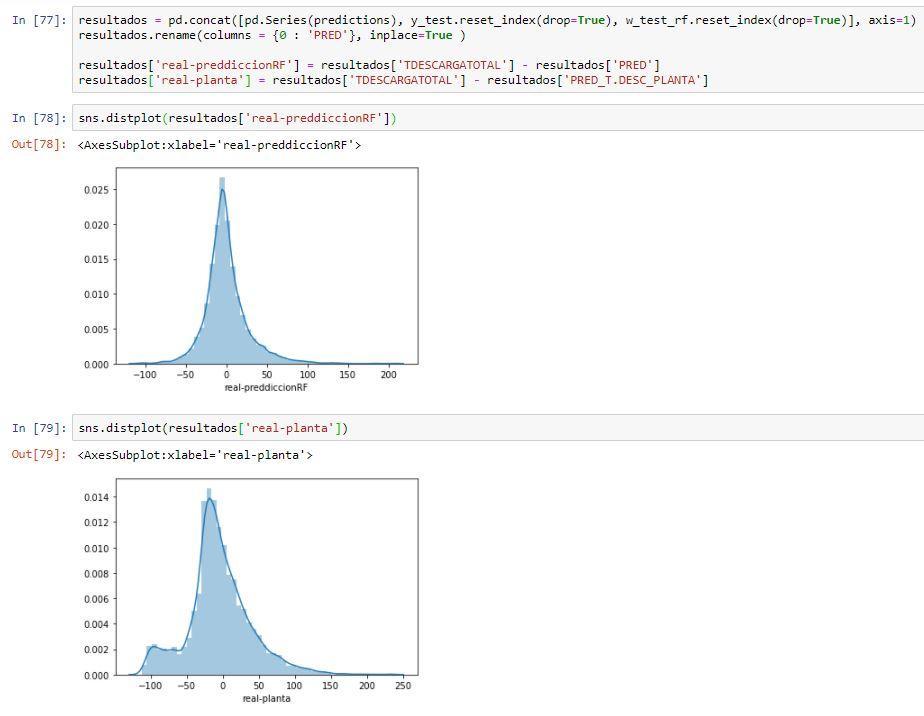


Se define como “X” la data con variables independientes, “Y” como la variable dependiente, y “W” como el valor usado en planta para planificar los despachos basados únicamente en la estructura.

A continuación, se presentan los resultados de cada modelo definido anteriormente:

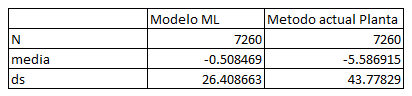
El modelo XGBRegressor con un rmse: 26,68 min es quien mejor performance para predecir tiene. Para hallar este número se ha usado la predicción X\_test vs Y\_test (dato real).

Adicionalmente, podemos visualizar el rmse de planta hallado a partir de Y\_test y W\_test equivalente a 44.3355 min. La distribución de los errores podemos visualizarla en la gráfica siguiente:



Como último paso se realizará una prueba de hipótesis de comparación de medias.

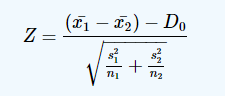
Se muestran las estadísticas de cada muestra.



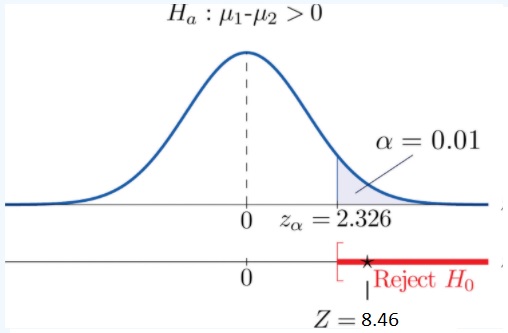
Hallamos el valor de Z, dado que las muestras contienen un n>30. Y usaremos un 99% de confianza, donde:

H0: X1 - X2 = 0

H1: X1 - X2 > 0



Según tablas, Z(99%) presenta un valor de 2.326, por lo tanto, se obtiene la siguiente gráfica



Se procede a rechazar H0, y se afirma con un 99% de confianza que X1 > X2.

Discusión de Resultados

En el siguiente apartado, incluimos las aportaciones en relación a los resultados obtenidos. Se ofrece un análisis exhaustivo relacionando los comentarios con las hipótesis iniciales.

Hay que partir de la base de que es fundamental que los datos con los que trabajamos sean fiables, de esta manera el resultado de nuestro modelo de predicción se ajustará más a la realidad. Para cumplir dicha premisa, dado que este proyecto se inició en abril de 2018, se ha tenido en cuenta un periodo de tiempo en el que ya se había implementado correctamente el procedimiento, dejando pasar los primeros meses de prueba y verificaciones de la información que se reporta. De esta forma, empezamos a trabajar los datos y a entrenar el modelo con información comprendida en el siguiente espacio de tiempo: Del 1 de agosto del 2019 al 31 de diciembre del 2020.

La información contenida en el periodo de tiempo indicado anteriormente, no tenía un formato y estructura uniforme, pues se tuvieron que hacer una serie de acciones para limpiar los 45.683 registros de despachos de hormigón de los que disponíamos de forma preliminar según lo explicado en el apartado resultados del presente trabajo

Concluidas estas acciones, tras la limpieza de registros erróneos, se cuenta con 36.298 registros aptos para poder entrenar el modelo. En este caso, tuvimos que definir las variables que entran en juego para poder entrenar el desarrollo con distintos modelos y ver los resultados. Sin perder de vista el objetivo de este proyecto, la mejora del tiempo de descarga, definimos ésta como la variable dependiente del proyecto ya que es la única parte del proceso que está dentro del control de la organización .

Dicha variable casi no correlaciona estadísticamente con el resto de variables independientes pero que inciden en los resultados de la predicción de tiempo de descarga que se quiere lograr. Para poder definir cuáles son las variables independientes más relevantes, se procede a revisar si se presenta alguna correlación significativa (>0.6) entre las variables independientes. El threshold de 0.6 se establece debido a la poca correlación encontrada en los análisis preliminares de las variables, teniendo en cuenta que estadísticamente hablando la convención es tomar los valores por sobre 0.8.

Tras el análisis de los resultados obtenidos, vemos que ninguna de las variables está expresamente correlacionada y que todas podrían funcionar en el modelo. No obstante, se decide conocer cuán importantes son, es decir, se pretende valorar la incidencia sobre la variable dependiente. Para ello, se procede a ponderar la importancia de cada una de las variables independientes mediante la creación de un modelo tipo “Random Forest Regressor”, entrenamos el modelo definido y visualizamos la relevancia de las mismas mediante el atributo “feature\_importances\_”.

Se escoge un total de 30 variables independientes de las 42 existentes ya que, dichas variables inciden en un 95% sobre la variable target que hemos definido anteriormente. Descartando de esta forma el % del total de variables descriptivas. Además, de las variables escogidas, se toma en consideración que las más relevantes son: la distancia, el volumen y la modalidad de despacho, pues la suma de estas nos indica el 40% de la correlación.

Esto contrasta en gran medida con la hipótesis inicial de la investigación, ya que en un principio la experiencia del negocio indicaba que el tipo de estructura a construir y la fórmula del hormigón serían muy determinantes en los tiempos de descarga. Sin embargo, después de haber realizado todos los análisis exploratorios dicha hipótesis descartada.

Tras todo lo anteriormente expuesto, el esquema queda definido de la siguiente forma:

Y\_Variable dependiente (Tiempo descarga)

x\_Variables independientes(Variables\_30)

w\_Actual (datos planificador actual).

En este punto, ya disponemos de todos los factores para comenzar a trabajar en hacer correr 4 modelos que se han tomado en consideración para verificar el modelo de predicción propuesto en este proyecto en comparación con el modelo actual existente en la compañía cementera.

|  |  |  |
| --- | --- | --- |
| Modelo | Resultado | Valor actual |
| Árbol de decisión | 28.77 min | 44.33 min |
| Random Forest | 26.76 min | 44.33 min |
| KNN | 29.37 min | 44.33 min |
| XGB | 26.68 min | 44.33 min |

Del análisis de resultados que se han reportado tras la aplicación de los 4 modelos definidos, se evidencia que el modelo que reporta mejores resultados sería el “modelo\_XGBRegressor” Ya que disminuye la variabilidad del tiempo de descarga en un % respecto a la actual. Esto es, aplicando el modelo la variabilidad sería de 26.68 minutos en comparación con los 44.33 minutos de variabilidad que hay a día de hoy.

Los hallazgos más importantes en términos de entendimiento del problema de negocio radican básicamente en que las variables que se intuía tendrían mayor repercusión en el tiempo de descarga, terminaron siendo secundarias para el estudio. Tal es el caso de los tipos de estructura, el volumen entregado y otras como los centros de distribución. Sin embargo, el componente distancia del dataset, al analizar a profundidad el problema agrega nuevos insights al negocio: a mayor distancia desde el centro de distribución a la obra, mayor tiempo de descarga. Y esto se puede interpretar en el hecho que mientras más tiempo está el hormigón en el mixer, va perdiendo su fluidez o trabajabilidad y ello conlleva a un mayor tiempo de descarga.

En cuanto al volumen a entregar y la modalidad, se comprueba claramente que a mayor volumen de entrega mayor tiempo de descarga, lo cual no representa ninguna sorpresa. De igual forma, al tener que incluir elementos adicionales para la descarga, como es el caso de una bomba, obviamente el proceso tiende a ser más lento.

Inmediatamente después, se observa que sí existen estructuras que determinan un mayor tiempo de descarga, sin embargo, el valor que aportan a la predicción del modelo es poco y por ello deberían pasar a un segundo nivel.

Para concluir si se logró el objetivo planteado, debemos tener un sustento estadístico que determine si un modelo es mejor que otro; se presentó una prueba de hipótesis de diferencia de medias sobre los errores de predicción, en donde se concluye con un nivel de confianza de 99%, que la media del modelo XGB es distinta y mayor que la del método actual. Adicionalmente se debe interpretar que la media con un valor más cercano a cero, es aquella que beneficiará a las predicciones al tener menos media de error.

Demostrado que el modelo desarrollado en la investigación nos ayudará a reducir el error de predicción, se puede afirmar que podemos hacer más eficiente el proceso de elaboración de programa diario, reduciendo la variabilidad del tiempo de descarga y por consiguiente el tiempo de ciclo, y así impactar en la satisfacción del cliente.

Conclusiones de gestión

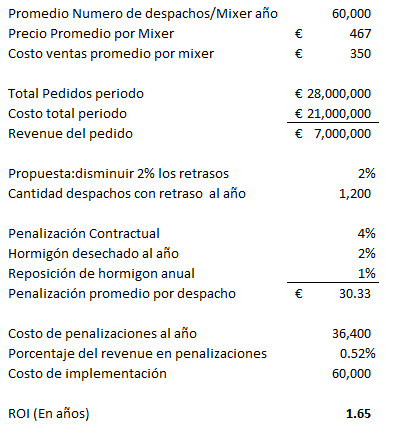
Fuera del análisis técnico propios a la investigación, de caras a la estructura organizativa también se puede rescatar información valiosa, la cual es importante implementar paralelamente con el modelo para asegurar un proceso fluido y el incremento de la satisfacción general de los clientes, ya que, si no hay una correcta coordinación interna de los procesos por parte de los departamentos involucrados, repercute en atrasos en los procesos de despacho. Para ello es indispensable la coordinación entre los dos departamentos que están implicados en el proceso, tomando en consideración las siguientes recomendaciones u objetivos adicionales, si se quiere:

* Mejorar la comunicación del departamento de ventas con el cliente: Establecer reuniones con los clientes de una periodicidad concreta (en función a los perfiles de cliente) para establecer relaciones de confianza y mantener las bases de datos con la información de los clientes actualizada.
* El técnico de ventas debe ofrecer una experiencia de venta mucho más completa en la que se informe de las características de todos aquellos servicios (modalidades de despacho) o productos (tipos de cemento) que disponemos en cartera y que mejor se ajustan al proyecto que quieren llevar a cabo. Para dar visibilidad al servicio que da la empresa hormigonera y que el cliente sienta seguridad y transparencia en la información que se le ofrece.
* Involucrar al departamento de programación en el momento que el técnico de ventas inicia la negociación del despacho: cuando se trata información relativa a estructura, coordenadas, modalidad de despacho… para que sea conocedor de las características del servicio para ese cliente en hormigón e intervenga para aportar valor si cree conveniente.

Para garantizar el correcto funcionamiento de estos objetivos, el responsable del departamento de ventas y el responsable del departamento de programación convocarán reuniones trimestrales de seguimiento para hacer un seguimiento de la puesta en marcha de estas acciones y evaluar la mejora que implican en la operativa. Además, se harán encuestas semestrales a los clientes para obtener feedback sobre dicha implementación y valoración del servicio.

Finalmente, para poder implementar la solución propuesta, el proyecto debe ser económicamente viable, por lo cual se presenta un análisis del ROI bajo los supuestos de mejorar de 10% a 8% el porcentaje de retrasos anuales.

Figura 9- Retorno de la Inversión



Reflexiones Finales

Durante la realización del presente trabajo de investigación se ha corroborado la importancia y relevancia de la optimización de los sistemas de programación de despachos para el caso específico de la empresa Hormigones Latinoamérica. Sin embargo, es importante destacar que las conclusiones acá expresadas son aplicables prácticamente a cualquier industria cuyo producto tenga que cumplir con tiempos de despacho acotados en un espacio de tiempo predefinido y cuyo incumplimiento tenga repercusiones tanto en la gestión de la organización como en la de sus usuarios finales.

En la medida que las herramientas de recolección, almacenamiento y análisis de los datos se democratizan y se hacen más accesibles a un mayor número de empresas, se puede observar como las buenas prácticas se replican tanto en el mismo ecosistema como en actividades económicas más alejadas del caso de estudio.

Precisamente es este el punto que se quiere rescatar del presente trabajo de investigación. Es nuestro entendimiento que prácticamente cualquier industria que dependa de un timeframe cerrado de entrega podría beneficiarse la aplicación de un modelo de predicción de tiempos de despacho. Sin embargo, es necesario profundizar en ciertas consideraciones para que este tipo de soluciones se conviertan en una metodología aplicable transversalmente y proviene precisamente de las dificultades encontradas y no siempre superadas durante el desarrollo de este caso. Entre ellas se pueden mencionar

* La calidad de la data recolectada es el punto más importante de cualquier implementación de este tipo. De nada vale contar con pentabytes de data almacenada si no se estandariza un proceso de ETL que sirva a los requerimientos actuales o futuros de la organización.
* Es fundamental que los sistemas de recolección cuenten con sistemas de comprobación redundantes durante el proceso de recolección de datos, y que las soluciones externas sean en realidad efectivas para los requerimientos de calidad de los datos.
* Es importante que los sistemas de recolección sean escalables. En el caso particular de esta investigación la opción de contar con sensores de telemetría abriría nuevos campos de investigación para modelos de predicción más avanzados como planeación de rutas, optimización de uso de combustible e incluso ampliación o disminución de la flota disponible con base a la demanda
* Los sistemas y modelos siempre pueden ser mejorados y se debe apuntar siempre hacia la mejora continua, incluso cuando los presupuestos de ID sugieran lo contrario, ya que el beneficio final de una solución bien implementada superará con creces la inversión.

A pesar de estas conclusiones generales también es importante destacar que, como todo sistema, no está exento de limitaciones técnicas y aplicativas, entre las cuales se pueden mencionar:

* La dependencia de la organización con entes externos para poder obtener más datos con mayor calidad. Esto se hace más claro al considerar que si quisiera aumentar la periodicidad del registro de señales de GPS de cada 60 segundos a cada segundo, por ejemplo, aumentaría exponencialmente el costo del sistema. Por esta razón es importante tener en cuenta hasta donde se quiere llegar.
* El aumento del costo de los datos y la tendencia de los fabricantes de hardware (incluyendo los vehículos) a vender la data como un adicional. A medida que aumenta la sofisticación de los sensores pre instalados en los sistemas y la capacidad de almacenarlos, cada vez más empresas están migrando a un sistema de comercialización de estos datos (PAAS o SAAS), y que inevitablemente se tendrá que pagar como un plus si se quiere aprovechar.
* Las integraciones intersistemas muchas veces son complejas y costosas. Para el caso de uso que atañe a esta investigación fue considerado incluir la variable clima dentro de la ecuación a través de una API externa, debido a la potencial influencia que dicha variable podría tener en los tiempos de despacho. Sin embargo, desde un punto de vista netamente académico no fue posible contar con los recursos necesarios para implementarlo en este modelo debido a su elevado coste (500 euros por el registro climático histórico por ciudad).
* Adicionalmente, existen limitaciones o dificultades específicas del producto, que requieren de la revisión y actualización constante de las reglas de negocio para poder garantizar una data consistente y que permita la alimentación automática de los modelos implementados. Un claro ejemplo de este problema es el hecho de que al utilizar sensores magnéticos y dado que la tierra contiene mineral ferroso, la misma crea campos magnéticos que pueden alterar el funcionamiento de los sensores, creando triggers “falsos positivos”, lo cual resulta en que la data de aproximadamente un 10% de los casos no sea fidedigna.

De cualquier manera y a pesar de las dificultades antes mencionadas, se entiende que todos los sistemas son siempre mejorables y escalables, el equipo de investigación del presente trabajo confía plenamente que los resultados expresados en él pueden ser fácilmente replicable en otras industrias cuyas entregas sean sensibles a ventanas cerradas de tiempo, como las que se mencionaron al inicio de la investigación.

Como próximos pasos y en aras de la mejora continua, se sugieren los siguientes próximos pasos para aumentar la eficiencia de los modelos y de los sistemas de recolección de datos:

* Elaboración de un nuevo proceso de encuesta tantos a los mismos clientes entrevistados en el origen de la investigación como a los diferentes departamentos internos de la organización que participan en el proceso, para corroborar que los resultados del modelo han aumentado el nivel de satisfacción, lo cual por motivos de tiempo y distancia quedaron fueran del alcance de este estudio.
* Inclusión de las variables meteorológicas de los sitios de descarga en el modelo, ya que existe la hipótesis de que el clima puede afectar directamente los tiempos de fraguado del hormigón y por ende su tiempo de descarga.
* Instalación de sistemas de telemetría e IOT para monitorizar el funcionamiento en conjunto del vehículo, para controlar variables mecánicas y pasar a un modelo de mantenimiento predictivo de las unidades, optimización de las rutas, el combustible y la seguridad de los conductores.
* En una etapa mucho más avanzada del proyecto, evaluar la inclusión de sistemas automatizados de conducción y descarga automática.

Referencias Bibliográficas

*(2008). Obtención del Cemento*. Recuperado de

<https://sjnavarro.files.wordpress.com/2008/09/def-y-elaboracion-cemento.pdf>)

Freitas, L., Borges, T., Martins, A., Panaino, R., Neto, J. y Filho, M.(2019). *Análise e proposta de redução de lead time no processo de corte, carregamento e transporte de uma usina de cana-de-açúcar no estado de São Paulo: um estudo de caso.* Recuperado de <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2019000300221&lang=es>

Hugas, J. (2019). *Cadena de suministro y logística: La transformación necesaria*. Recuperado de

<https://www.harvard-deusto.com/cadena-de-suministro-y-logistica-la-transformacion-necesaria>

Rodriguez, E. (2020)*. El puente hacia una logística inteligente.* Recuperado de

<https://www.harvard-deusto.com/el-puente-hacia-una-logistica-inteligente>

Lehmann, E. L.; Casella, George (1998). Theory of Point Estimation (2nd edición). Nueva York: Springer. ISBN 0-387-98502-6. MR 1639875.

Sheffi, Y., Saenz, M. J., Rivera, L., & Gligor, D. (2019). New forms of partnership: The role of logistics clusters in facilitating horizontal collaboration mechanism. European Planning Studies, in press

Bot, B. y Neumann, C. (2003).*Hacia la optimización de la cadena logística.* Recuperado de

<https://www.harvard-deusto.com/hacia-la-optimizacion-de-la-cadena-logistica>

*Riaño, L. (2006). La logística y las operaciones como factores críticos de éxito.* Recuperado de

<https://www.harvard-deusto.com/la-logistica-y-las-operaciones-como-factores-criticos-de-exito>

B. Gesing, S. J. Peterson, D. Michelsen, Artificial Intelligence in Logistic; A collaborative report by DHL and IBM on implications and use cases for the logistics industry, page 15, 2018.(Gesing, Peterson, & Michelsen, 2018

Viuda. J (2006). *La excelencia en la cadena de suministro : ¿Fuente de ventaja competitiva?.* Recuperado de

<https://www.harvard-deusto.com/la-excelencia-en-la-cadena-de-suministro-fuente-de-ventaja-competitiva>

Leporati, M. y Contreras, M. (2019)*. Inteligencia artificial en la gestión de cadenas de suministro*. Recuperado de

<https://www.harvard-deusto.com/inteligencia-artificial-en-la-gestion-de-cadenas-de-suministro>

M. Breunig, R. Kelly, R. Mathis, D. Wee, Getting the most out of industry 4.0 (April 2016). Recuperado de URL <http://www.mckinsey.com/business-functions/operations/our-insights/industry-40-looking-beyond-the-initial-hype>

Ceballos Arana, Martín A. EL CONCRETO, MATERIAL FUNDAMENTAL PARA LA INFRAESTRUCTURA. Recuperado de http://www.revistacyt.com.mx/pdf/agosto2016/experto.pdf

Jordisolivir (2020), *¿Que es el Transporte neumático de Sólidos?*. Recuperado de

<https://solivir.es/que-es-el-transporte-neumatico-de-solidos/>

ENCUESTA DE SATISFACCIÓN HORMIGONES LATINOAMÉRICA C.A , (2016).

# 

# Anexos

Anexo 1. Ejemplo de Reporte de Programación

# 

Anexo 2: Tecnología de sensores unidades e instalación:

IOT HUB



3G/2G (GSM/GPRS/SMS), BLE, Resistente al agua: (IP65)

* Batería de respaldo: 250mAh (Max 2 Hrs)
* Almacenamiento: 30,000 eventos
* Acelerómetro: 3-Axis Digital ±16G
* Interface: RS232, USB, 1-Wire, ADC, I/Os, JS ON/APIs (OTA), Bluetooth (BLE)
* IoT Ready: (JSON/APIs OTA interconnectivity support



* Unidad de control de motor.
* Interface para leer protocolos CAN J19 39 / J1708 & J1587 para aplicaciones vehiculares.



* Unidad de monitoreo conductiva de tambor de mezcla y su instalación





Anexo 3

Notebook conteniendo los códigos de la investigación: (Ver PDF Anexo)