Modelo de Estimación de Oferta y Demanda de Productos usando Series de Tiempo y Algoritmo de Clasificación de K-Vecino más Cercano

Alfredo José Naime Marcano
Dirección de Automatización, Informática y Telecomunicaciones (AIT)
Petróleos de Venezuela S.A. (PDVSA)
Caracas, Venezuela
naimea@pdvsa.com

***Resumen*—Este trabajo busca construir modelos predictivos para la estimación de volúmenes de oferta y demanda de productos, usando modelos de Series Temporales y Algoritmo de K-Vecinos más Cercanos (KNN). Los modelos preseleccionados por sus atributos estadísticos, son clasificados usando KNN para determinar su calidad en "Excelentes", "Buenos" y "Malos". Cuando la clasificación es dicotómica, obtenemos un 67,75% de acierto para modelos "Buenos" y un 57,67% de acierto para modelos "Malos". Dentro del grupo de los tres modelos mejor clasificados, que se le presentan al decisor, el porcentaje de acierto llega al 75%. Para una de las señales estudiadas, en los 22 períodos observados se encontró un modelo "Excelente", de los cuales 19 (86,36%) de ellos tuvieron ese mismo resultado y 3 tuvieron un resultado dentro del rango "Bueno".**

Palabras Clave*—*Modelos Predictivos, Series de Tiempo, Minería de Datos, Aprendizaje de Máquina, k-vecinos más Cercano, Modelos de Oferta y Demanda, Estimación, Pronósticos.

I. Introducción

Los modelos de oferta y demanda de productos buscan describir el comportamiento de la producción y el consumo de bienes, principalmente para buscar el equilibrio del mercado y la estabilización del precio del producto. Marcelo Resico [1], indica que "la actividad humana en relación con los medios se torna económica en cuanto estos medios son escasos. Y, evidentemente, si los medios que necesitamos fueran ilimitados no habría necesidad de administrarlos de manera económica o economizarlos".

Más adelante, define el consumo como "la actividad humana dirigida a satisfacer las necesidades y deseos de las personas" y la producción como la actividad de donde derivan los bienes y servicios. Igualmente, menciona que "la distribución no sólo depende de factores evaluados por el propio sistema económico, sino también de la equidad y justicia social, que abarca los resultados del mercado como un elemento de la ética social" [1].

Cuando la demanda aumenta y la oferta permanece constante o disminuye, habrá escasez y esto forzará un aumento del precio del producto. Contrariamente, si la oferta aumenta y la demanda permanece constante o disminuye, habrá superávit del producto y el precio tenderá a bajar (Figura 1). Si bien, el precio es un elemento de gran relevancia tanto para productores como consumidores, grandes desequilibrios de estos factores traen inconvenientes adicionales a considerar, como por ejemplo, sobrepasar la capacidad de almacenamiento, transporte o distribución, producción diferida, pérdida de empleos, cambios en la matriz de importación/exportación, entre otras.

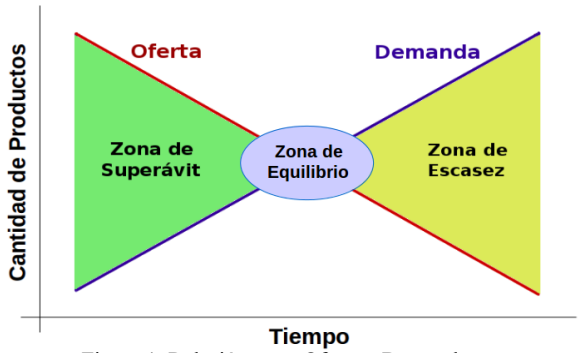


Figura 1. Relación entre Oferta y Demanda

Por tanto, mantener un balance de la oferta y la demanda de un producto y pronosticar el comportamiento del sistema es una habilidad muy deseada por economistas, gerentes de operaciones, administradores, inversionistas y gobierno, llegando a nivel de asunto de Estado en los sectores más influyentes de la economía y la industria de un país.

Nuestro trabajo se basa en construir modelos predictivos para la estimación de oferta y demanda de productos o bienes, privilegiando los volúmenes en lugar de los precios, como apoyo a la planificación logística y operacional, sin embargo, puede servir como insumo para establecer modelos econométricos a partir de los volúmenes estimados o incorporarse en un modelo integral. Dentro de los mercados de bienes, en una visión microeconómica [1], podemos aplicar estos modelos a bienes de consumo final e intermedios o insumos, tales como madera, piezas, electricidad o combustible [1].

En la siguiente sección describiremos el sistema a modelar y cómo se calcula su balance. Posteriormente, planteamos la arquitectura de solución, seguido de algunos detalles de implantación con respecto a los ETLs (programas de extracción, transformación y carga de datos) y la Base de Datos. En la Sección VI hablaremos de los aspectos

relacionados con la Minería de Datos, cuyos resultados estarán planteados en la Sección VII. El documento finaliza con las secciones de conclusiones, recomendaciones y referencias.

II. Descripción del Sistema y su Balance

En un sistema de oferta y demanda, tendremos por un lado a productores, que aportan un producto de determinado tipo (camisas, alimentos, líquidos, gas, etc.), que se puede cuantificar en alguna unidad de medida (unidades, docenas, toneladas, litros, pies cúbicos, etc), y por el otro lado, tendremos clientes o usuarios que los consumen (consumo expresado en las mismas unidades de la oferta). Entre estos dos grupos existe una red de distribución o mercado que los interconecta (Figura 2).

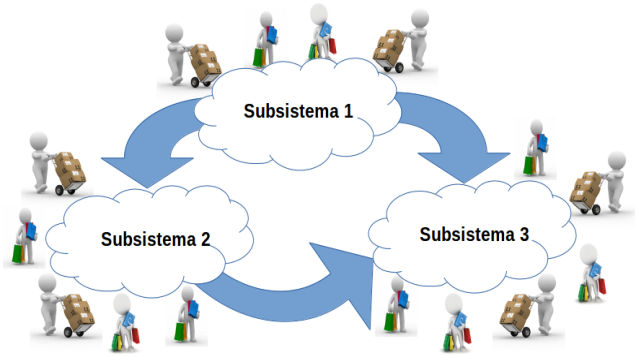


Figura 2. Sistema de Oferta y Demanda

Esta red puede verse a distintos niveles de detalle y los productores/consumidores pueden asociarse por distintos criterios, como geográficos (región, provincia, municipio), sector económico, industrial o comercial, tipo de organización u otros, pudiendo crear subsistemas separados física o lógicamente y puntos de transferencia o contacto entre subsistemas, por donde pasan productos de una subred a la otra.

Balance de Producto A (kgs)

Fecha 31/05/2018 6:00 am Ver decimales PDF

	Recibido	Entregado	Balance
Total Nación	2.135,83	2.077,04	58,79

Balance Regiones

	Recibido	Entregado	Balance
Total Oriente	1.686,35	1.615,20	71,15
Total Occidente	500,48	512,83	-12,36

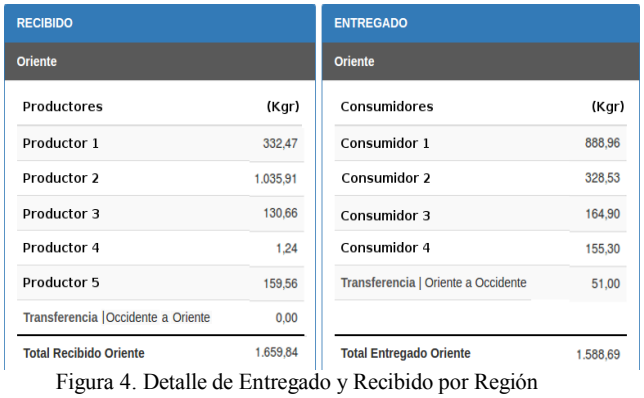
Transferencias

	Occidente a Oriente	Oriente a Occidente	Total
Transferencias	0,00	51,00	51,00

Figura 3. Balance Global y Regional

A nivel general y por subsistema se calcula un balance diario (Figura 3), el cual puede quedar en déficit, superávit o estable (oferta igual a demanda). El superávit se almacena o "empaca", por lo que la capacidad almacenada puede compensar el déficit por un tiempo determinado. Cada valor diario, puede estar en alza, baja o estable, en comparación con el día anterior y con estos indicadores se realiza el primer análisis del balance en este sistema.

Por cada subsistema, por los que se puedan agrupar a los usuarios o clientes y los proveedores, se puede calcular un balance entre oferta y demanda (Figura 4), pero se debe tener un especial cuidado al tratar las transferencias entre subsistemas. Una transferencia, será un consumidor o cliente en el subsistema de salida y un productor en el subsistema de entrada, pero estas entradas y salidas no aportan al total nación o global, ya que simplemente se mueven dentro del sistema global.



RECIBIDO — Oriente

Productores	(Kgr)
Productor 1	332,47
Productor 2	1.035,91
Productor 3	130,66
Productor 4	1,24
Productor 5	159,56
Transferencia |Occidente a Oriente	0,00
Total Recibido Oriente	1.659,84

ENTREGADO — Oriente

Consumidores	(Kgr)
Consumidor 1	888,96
Consumidor 2	328,53
Consumidor 3	164,90
Consumidor 4	155,30
Transferencia | Oriente a Occidente	51,00
Total Entregado Oriente	1.588,69

Figura 4. Detalle de Entregado y Recibido por Región

El Total Recibido Global (TRG) es la suma de lo aportado por cada productor, pero cuando se trata de un subsistema, a la suma de los productores asociados al mismo, se debe agregar las transferencias que provengan de otros sistemas. De forma análoga se calcula el Total Entregado Global (TEG), como la suma de lo consumido por cada cliente. Cuando el cálculo de los totales globales (TRG o TEG), se realiza a partir de los regionales o subsistemas (TRR y TER), se deben restar las transferencias, las cuales son puntos donde pasan mercancías de un subsistema a otro, extrayendo/aportando al subsistema pero no al cálculo global. El Balance Global (BG) es lo recibido por la red menos lo entregado.

$$TRG=\sum_{i=1}^{N_p} p_i=\sum_{j=1}^{N_r} TRR_j-\sum_{k=1}^{N_t} Trans_k$$

$$TEG=\sum_{i=1}^{N_c} c_i=\sum_{j=1}^{N_r} TER_j-\sum_{k=1}^{N_t} Trans_k$$

$$BG=TRG-TEG=\sum_{i=1}^{N_r} TRR_i-TER_i$$

Donde :
N_p : *Número de productores*
N_c : *Número de consumidores*
N_r : *Número de regiones o subsistemas*

Por cada área, región o subsistema y serie registrada (productor o consumidor), se calculan estadísticas de producción y consumo. Estos datos, analizados en el tiempo a intervalos regulares (por lo general diario), pueden ser vistos como series temporales con los que podemos hacer tanto análisis descriptivo como predictivo. En la Figura 5, una

vista de los estadísticos descriptivos en forma de diagrama de cajas (*boxplots*), mostrando la distribución de los datos para el período semanal, mensual y anual anterior.

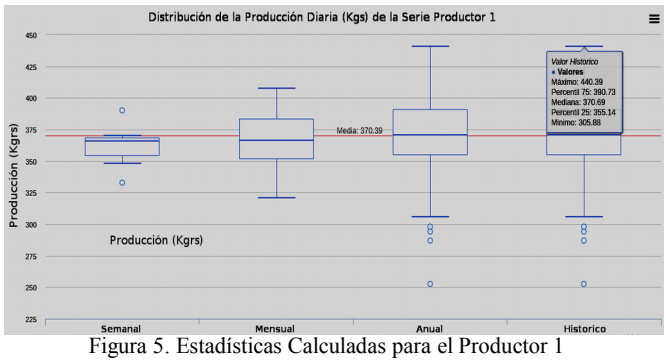


Figura 5. Estadísticas Calculadas para el Productor 1

Finalmente, para cada serie se realiza un proceso de cálculo intensivo para la estimación por series de tiempo [2] de los siguientes siete períodos, basado en los cien períodos anteriores (Figura 6). Mediante un proceso de clasificación, usando el algoritmo de K-vecino más cercano, se seleccionan los tres mejores modelos para hacer el pronóstico. En las siguientes secciones, describiremos la metodología y diseño del sistema y los resultados obtenidos con el algoritmo de K-vecino más cercano a un mes de su puesta en producción. Debido a la confidencialidad de los datos operacionales que maneja este sistema, los resultados serán descritos usando medidas referenciales. Nuestro modelo, para el momento de este escrito, consiste de once señales de nodos productores y ocho nodos consumidores, distribuidos en dos regiones (Oriente y Occidente).

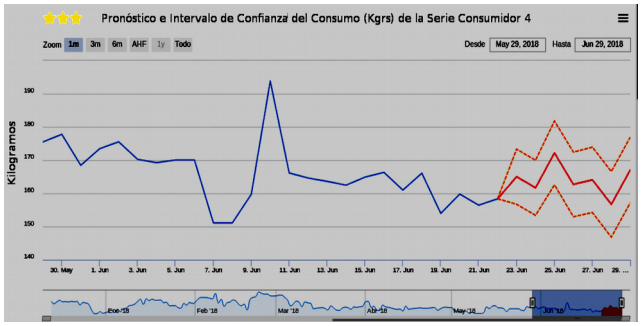


Figura 6. Serie de Datos con Pronóstico del Consumidor 4

III. Arquitectura de Solución

La arquitectura de solución consiste en un conjunto de componentes que interactúan entre sí para extraer, transformar y validar los datos operacionales desde los sistemas de control o gestión empresarial, almacenándolos en una base de datos preparada para el análisis. Posteriormente se procesan y calculan las estadísticas y modelos predictivos, para finalmente clasificar los modelos usando algoritmos de aprendizaje de máquina. Los resultados de cada fase, se almacenan en la base de datos para su despliegue y consulta en línea por parte de los usuarios y del gerente operacional a medida que se van generando los mismos (Figura 7).

La extracción, transformación y caga de los datos se realiza con *Pentaho Data Integrations* y *Python*, la Base de Datos principal es *PostgreSQL*, los cálculos de estadísticas y los modelos predictivos se realizan con *R Studio Server* y la interfaz de usuario es elaborada con *PHP* y *JavaScript* usando *Laravel* con *Bootstrap*, *HighCharts* y *Jquery*, entre otras librerías.

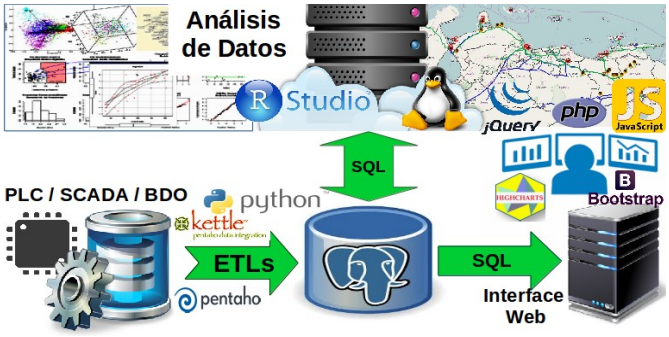


Figura 7. Arquitectura de Solución

En este escrito mostraremos sólo los componentes de Base de Datos del modelo operacional y los aspectos necesarios para el cálculo del balance y los modelos de minería de datos asociados a éste. El resto de los modelos dimensionales para Inteligencia de Negocios, los *data marts* y la evolución hacia el *data lake*, son temas que se plantearán en otros escritos.

IV. ETLs y Secuenciamiento

A través de los programas de Extracción, Transformación y Carga (ETL), se obtienen los valores de consumo y producción diarios de cada uno de los entes del sistema, ya sean como señales agrupadas o desagrupadas, según las condiciones del proceso a modelar. Las señales pueden obtenerse individualmente, de cada fuente o consolidadas, dependiendo de las condiciones operacionales, de seguridad de los datos y de los acuerdos entre las partes.

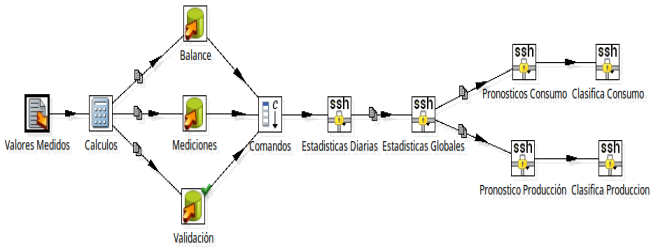


Figura 8. ETL para Carga y Secuenciamiento

De igual forma, estas tareas (*jobs*) programadas periódicamente, nos pueden servir para disparar actividades y procesos en distintos ambientes. El secuenciamiento de tareas nos permite no sólo realizar la manipulación de los datos, sino ejecutar actividades de análisis, cómputo científico, despliegue de resultados y prácticamente cualquier actividad que pueda realizar una computadora en una red.

La transformación de la Figura 8, describe una carga de datos desde un archivo fuente. Los mismos pasan a un proceso donde se realizan cálculos a partir de estos datos, los cuales pueden incluir, reconversión de unidades, agrupamiento o desagrupamiento por diversas reglas, detección y corrección de errores, cambios de formato, entre otras. De allí, los datos son divididos y cargados en dos tablas distintas, según su entorno y una que se usa para validación y

comprobación del proceso. Después se cargan unos comandos de sistema que serán ejecutados posteriormente en la nube de cómputo vía SSH, como se detalla en los siguientes componentes del secuenciamiento: se calculan estadísticas, se ejecuta el algoritmo de pronóstico y por último con estos resultados se clasifican los modelos que seleccionó el primer estimador.

Adicionalmente, los datos de producción y consumo diario deben ser comparados con el pronóstico realizado el día anterior para determinar cuan cerca o lejos estuvo el pronóstico del valor real (Figura 9). Para esto, con los mismos datos recibidos del día corriente, se ejecuta un proceso ETL que compara los pronósticos del día anterior con el valor actual por cada señal y las diferencias y los indicadores de eficiencia y precisión de cada modelo son guardados en otra tabla para evaluar las bondades de los algoritmos de aprendizaje automático implementados y plantear su evolución y mejora continua.

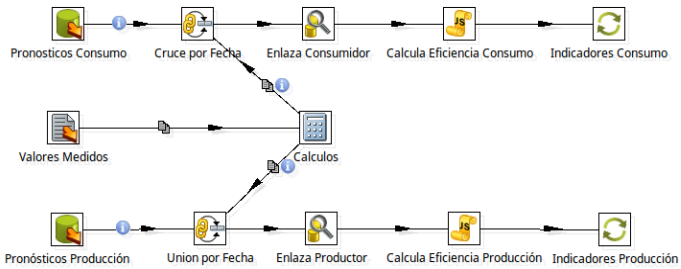


Figura 9. Actualización de Indicadores de Precisión

V. Base de Datos

La Base de Datos para esta solución está dividida en dos esquemas, el primero, netamente de registro operacional, con las señales que vienen de las distintas fuentes de datos y el segundo, para el análisis estadístico y predictivo, donde se almacenan los pronósticos que se calculan diariamente. Si se requieren datos por turnos o a distintos momentos del día, mientras la medición sea a intervalos regulares y fijos, sólo hay que agregar el componente de hora como un campo adicional. Sin embargo, si los intervalos no son regulares ni fijos, los datos registrados no podrán ser tratados en forma natural como una serie temporal. En este caso, la recomendación es trabajar con el indicador diario que mejor se ajuste al proceso y a los objetivos del problema, ya sea el valor total del día, el promedio, el máximo o el valor puntual a una determinada hora.

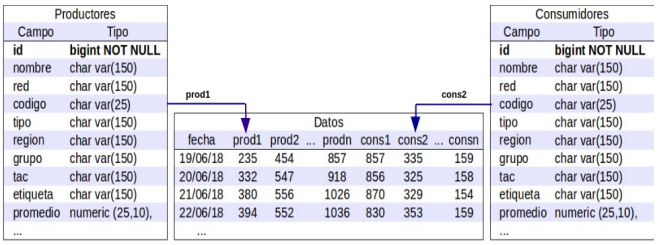


Figura 10. Esquema de Referencia entre Tablas

Para el primer ambiente, el modelo es bastante simple, los datos de consumidores y productores son almacenados cada uno en una tabla con un identificador que actúa como clave primaria. Los datos de consumo y producción de cada día son almacenados en una tabla con fecha como clave primaria y una columna para cada señal. Dentro de los datos de cada proveedor y cada consumidor, se registra un atributo para enlazar con la columna correspondiente a su serie o señal en el tiempo (Figura 10). Incluso, de considerar mantener señales agrupadas, totales y subtotales, para ser analizados como series temporales, se mantienen en la misma tabla de datos por fecha y los atributos de agrupamiento, clasificación y toda la metadata asociada al ente de producción o consumo se maneja en la tabla correspondiente.

Las estadísticas y cálculos de la distribución semanal, mensual, anual e histórica, se realiza a partir del análisis temporal de cada columna y son almacenados por fecha y por señal cuatro vectores con los estadísticos principales de tendencia central, dispersión y los datos atípicos (*outlier*) encontrados. Los cuatro vectores corresponden al análisis de los últimos 7, 30 y 365 días de la serie, completando con el cálculo de los mismos valores para la serie completa. Cada vector tendrá mínimo cinco valores: mínimo, percentil 25, percentil 50, percentil 75 y máximo. A partir de este punto, se colocarán todos los valores atípicos encontrados para esta serie, si existen (Figura 11).

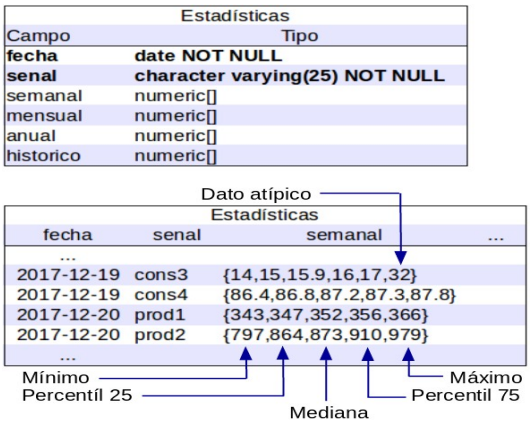


Figura 11. Estructura de Datos para Estadísticas

VI. Minería de Datos

Este proceso, es largo y con muchas posibilidades. Desde aplicar distintos algoritmos de clasificación y aprendizaje profundo (*deep learning*), hasta integrar otras capas de datos e información con tecnologías de *Big Data* y *Data Lake*. Nuestro prototipo inicial y base para este proyecto contempló la clasificación de los modelos con árboles de decisión [2] y en esta primera implantación en producción estamos usando el Algoritmo de K-Vecinos más Cercanos (KNN) [3][4][5][6].

Como todo proceso de Minería de Datos, comenzamos a preparar la base minable, la cual consistirá de todos los pronósticos realizados anteriormente por el algoritmo de pronóstico. Recordemos que este algoritmo [2], calcula alrededor de unos 1.152 modelos *ARIMA (p,d,q)* por cada serie de datos y basado en algunos indicadores estadísticos y pruebas de normalidad de residuos, selecciona los 16 modelos que mejor ajustan los datos. Entre estos indicadores podemos mencionar, el *Criterio de Información de Akaike* (AIC), el *test de Box-Pierce* y lo cerca que el modelo llega a predecir el valor *n* (último valor conocido) con los *n-1* valores anteriores, entre otros criterios. Estos indicadores

serán normalizados para su mejor interpretación por el algoritmo, al comparar las distancias entre vecinos.

A esto se le agrega la variable de clasificación, que indica la calidad del modelo como "Excelente" para los que presentan una diferencia menor a un 1% del valor reportado de producción o consumo, "Bueno" para los modelos cuyo pronóstico presenta una diferencia menor a un 5% del valor reportado y "Malo" para los demás, cuya diferencia estuvo por encima del 5%. Todos estos indicadores son consultados de la base de datos una vez se concluye el pronóstico para el día siguiente. Evidentemente, todos los pronósticos anteriores tienen una clasificación conocida porque ya el valor reportado para esos días se conoce (este es nuestro grupo de entrenamiento), pero el pronóstico realizado en ese momento para el día siguiente no tiene clasificación aún (este será nuestro grupo de prueba). Son estos modelos los que trataremos de clasificar a priori con el algoritmo de KNN, basado en lo cerca o no que quedaron los anteriores pronósticos cuyo resultado es conocido.

Con esta clasificación a-priori, se recalcula la valoración de los modelos, tomando en cuenta el resultado que dio el algoritmo de clasificación y el orden que traían de la primera fase, dada por el algoritmo de estimación por series temporales. De esta forma, los modelos clasificados como "Excelente", tendrán una valoración mayor que los clasificados como "Bueno" y éstos también por encima de los clasificados como "Malo". Dentro de cada grupo, se ordenan por los criterios del primer algoritmo. Al usuario, se le presenta la calidad del modelo como un componente gráfico representado con tres, dos y una estrella respectivamente y se le dan las tres mejores opciones para su estimación (ver Figura 6). Una vez conocido el resultado, se compara la clasificación a priori del algoritmo con el resultado real y se evalúa el error de clasificación. En esta primera fase, probamos clasificar con un k igual a 69, por ser la raíz cuadrada de la cardinalidad del conjunto de aprendizaje.

VII. Resultados

En la Tabla I, se puede ver la comparación entre la clasificación que realizó el algoritmo y el resultado que obtuvo después de conocido el valor real reportado para todos los modelos evaluados tanto para producción como consumo desde el 05 al 26 de Junio de 2018.

TABLA I. Clasificación de Modelos Predictivos

		Resultado			
		Excelente	Bueno	Malo	Total
Clasificado	Excelente	250	185	82	517
	Bueno	423	764	690	1.877
	Malo	190	743	1.271	2.204
	Total	863	1.692	2.043	4.598

TABLA II. Porcentaje de Acierto por Grupo de Clasificación

		Resultado			
		Excelente	Bueno	Malo	Total
Clasificado	Excelente	48,36%	35,78%	15,86%	11,24%
	Bueno	22,54%	40,70%	36,76%	40,82%
	Malo	8,62%	33,71%	57,67%	47,93%
	Total	18,77%	36,80%	44,43%	100,00%

Aquí podemos ver, en la diagonal, que el algoritmo acertó en la mayoría de los casos. De los 4.598 modelos en total, 517 (11,24%) fueron clasificados como excelentes (ver Tabla II) y de éstos 250 (48,36%) resultaron realmente excelentes en su estimación (es decir, llegaron a un pronóstico cuyo valor estuvo a menos de un 1% del valor reportado). De los 1.877 (40,82%) clasificados como buenos, 764 (40,7%) resultaron realmente modelos buenos, cuyo pronóstico estuvo a menos de 5% del valor reportado. Sin embargo, 423 (22,54%) resultaron por encima de lo esperado y tuvieron un resultado excelente. De los 2.204 (47,93%) modelos clasificados como "Malos", 1.271 (57,67%) coincidieron con este diagnóstico.

Ahora, si viéramos sólo dos grupos de clasificación y sólo distinguiéramos entre "Buenos" y "Malos". En este caso, tendríamos un 67,75% (ver Tabla III) de acierto para los modelos "Buenos" y un 57,67% de acierto para los modelos "Malos". Se pudiera pensar que el primer error (clasificar modelos "Malos" como "Buenos" es peor que el segundo (clasificar modelos "Buenos" como "Malos"), pero recordemos que los modelos "Malos" serán descartados, si existe algún otro clasificado como "Bueno" o "Excelente".

TABLA III. Porcentaje de Acierto con Clasificación Dicotómica

		Resultado		
		Bueno	Malo	Total
Clasificado	Bueno	67,75%	32,25%	52,07%
	Malo	42,33%	57,67%	47,93%
	Total	55,57%	44,43%	100,00%

Por tanto, evaluemos el porcentaje de acierto, dentro del grupo seleccionado para estimar, es decir, dentro de los tres modelos que se le presentan al decisor para que realice la estimación de la siguiente semana de producción o de consumo. En la Tabla IV, podemos observar como el 69,26% de los modelos presentados fueron clasificados como "Buenos" o "Excelentes" (aquí agrupados), de los cuales tres de cada cuatro (75%) cumplieron con esa clasificación, mejorando un poco más de 7 puntos porcentuales. En los períodos en los que no hubo modelos con mejor clasificación y por tanto, se presentaron al decisor modelos clasificados como "Malos", el porcentaje de error se mantiene muy similar al anterior. Sin embargo, como es de esperar, estos modelos representan menos de la tercera parte, ya que sólo son seleccionados cuando no existe una mejor opción según la valoración otorgada por el algoritmo.

TABLA IV. Tasa de Acierto en Modelos para Estimación

		Resultado		
		Bueno	Malo	Total
Clasificado	Bueno	75,00 %	25,00 %	69,26 %
	Malo	41,55 %	58,45 %	30,74 %
	Total	64,72 %	35,28 %	100,00 %

Finalmente, es de esperar que el algoritmo no sea igual de eficiente con todas las señales de consumo o producción evaluadas, a continuación veremos uno de los casos de producción y de consumo, tomando la mejor predicción seleccionada para cada periodo y como el pronóstico se fue acercando al valor real reportado durante el periodo observado. Para esta señal, los 22 períodos observados tuvieron como primera opción un modelo clasificado excelente, de los cuales, en 19 oportunidades o períodos (86,36%) el modelo seleccionado cumplió con ese mismo

resultado y en 3 períodos su resultado estuvo dentro del rango bueno. En la Figura 12, se aprecia la gráfica comparativa del pronóstico y el valor real reportado.

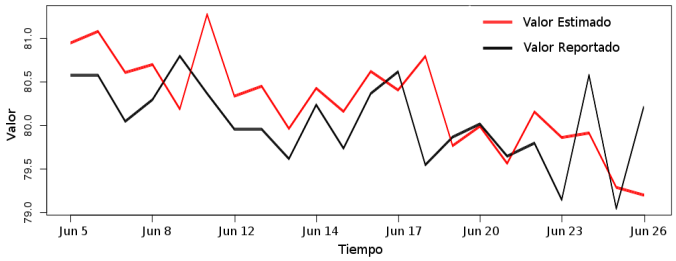


Figura 12. Pronóstico vs Valor Reportado (Productor 11)

Para la serie del Consumidor 7 (Figura 13), como primera opción dada al decisor, en 4 períodos se encontraron modelos clasificados como excelente y en 18 períodos modelos clasificados como buenos. De los 4 clasificados excelentes, tres tuvieron ese mismo resultado y uno de ellos tuvo un resultado bueno. De los 18 clasificados buenos, 6 dieron resultado excelente, 10 coincidieron en el resultado y sólo dos no cumplieron las expectativas, resultando malos pronosticadores.

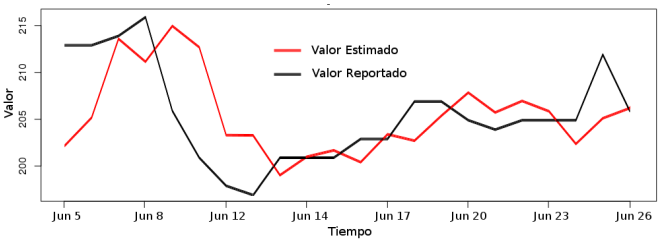


Figura 13. Pronóstico vs Valor Reportado (Consumidor 7)

VIII. Conclusiones

Tratar de predecir el futuro es una de las actividades más difíciles y arriesgadas, cuando aciertas, muchos piensan que era evidente y cuando te equivocas también. Sin embargo, siempre es natural tener una "mejor visión" del futuro, ya sea con base en la evidencia y los datos disponibles o en la pura intuición. El uso de modelos de series temporales, combinados con distintos algoritmos de aprendizaje estadístico puede y seguirá dando aportes para mejorar nuestros pronósticos, tanto en el sector industrial como comercial.

Para esta primera implantación, el algoritmo de k-vecinos más cercanos, no solo dio muy buenos resultados y cumplió con los objetivos propuestos, sino que resultó un método simple, rápido y eficaz de clasificación en este escenario. Pudiera plantearse mejorar la tasa de aciertos, tanto segmentando la muestra de aprendizaje como adaptando los parámetros y algoritmos dependiendo de cada señal analizada.

La evolución de estas prácticas puede ir en muchos sentidos e integrarse con aún más posibilidades de análisis, comparando estos resultados y pronósticos operacionales con los comerciales y estratégicos, a nivel local, regional y nacional. El predecir los niveles de escasez o superávit en un determinado producto en un mercado determinado, ayuda tanto a la planificación operacional y comercial, como a la planificación estratégica de un sector y de un país.

Mejorando la calidad y oportunidad de los datos e indicadores de producción y consumo para un determinado producto estratégico, es posible construir modelos dinámicos que apoyen la toma de decisiones estratégicas y operacionales de empresas públicas y privadas, así como para la planificación de políticas públicas y macroeconómicas más ajustadas a la realidad.

IX. Recomendaciones

Recomendamos seguir explorando alternativas dentro de los algoritmos de aprendizaje estadístico y *deep learning*, así como la integración con otras fuentes de datos, para maximizar el potencial de estos algoritmos en el tiempo, a medida que obtenemos mayor conocimiento de los procesos modelados y de los factores que los afectan, así como la expectativa o "sentimiento" del mercado, el cual es un aspecto esencial en todo modelo donde interactúa el ser humano y donde sus necesidades, miedos y aspiraciones, juegan un papel importante.

Integrar estos modelos de predicción con variables econométricas puede apoyar con mayor información a los decisores comerciales y estratégicos, así como a mejorar los modelos de valoración económica de las mercancías y los recursos estratégicos de un país.

Integrar los modelos operacionales, de corto plazo, con modelos tácticos y estratégicos, de mediano y largo plazo, puede brindar posibilidades mucho mayores de análisis y de optimización en la gestión de recursos en diversos sectores de la vida nacional.

Referencias

[1] M. F. Resico. "Introducción a la Economía Social de Mercado". Konrad Adenauer Stiftung. Argentina. Noviembre 2010. pp. 38-67. ISBN: 978-987-1285-21-1.

[2] A. J. Naime. "Algoritmo Adaptativo para la Estimación por Series de Tiempo Usando Arboles de Decisión". Quinta Conferencia Nacional de Computación, Informática y Sistemas / CoNCISa 2017. pp. 134-140. 2017. ISBN: 978-980-7683-03-6.

[3] W. N. Venables and B. D. Ripley. Modern Applied Statistics with S. Fourth edition. Springer 2002. pp. 350-358.

[4] P. K. Janert. Data Analysis with Open Source Tools. O'Reilly 2010. pp. 428-451. ISBN: 978-0-596-80235-6.

[5] I. H. Witten and E. Frank. Data Mining, Practical Machine Learning Tools and Techniques. Morgan Kaufmann Publishers, Elsevier. Second Edition. 2005. pp. 76-80, pp. 128-254.

[6] N. Zumel and J. Mount. Practical Data Science with R. Hanning Shelter Island. pp. 156-166. ISBN: 9781617291562.