

Digital **NAO**

Machine Learning. Algoritmos y modelos de predicción



PRIMER CONTACTO

Actividad

Hablar sobre movilidad en Corea del Sur no es algo sencillo. La concentración de población en las ciudades importantes de este país provoca que se trate con cuidado este tema debido a su naturaleza compleja.

Por eso, en el 2015, el Ministerio de Movilidad de Corea del Sur impulsó diversas iniciativas con el fin de complementar las capacidades de la infraestructura de transporte en las principales ciudades, no solo desde una perspectiva sustentable con el medio ambiente, sino también para buscar un aprovechamiento completo de las inversiones.

En este contexto, la ministro de Movilidad, Soo Yeon, inauguró en el año 2017 en la ciudad de Seúl el Sistema BikePro, el cual permite a los ciudadanos registrarse para que, por medio de un sistema online de alquiler, puedan viajar en una bicicleta de cualquiera de sus estaciones situadas alrededor de la ciudad, y devolverla en la estación más cercana a su destino, considerando trayectos que no excedan un par de horas.

BikePro fue muy bien recibido entre el público pues ofrece esquemas de renta muy flexibles y económicos, las estaciones están bien ubicadas en puntos de importancia urbana, como campos estudiantiles, oficinas de trabajo y centros comerciales. Además, su horario de funcionamiento, es otra ventaja, ya que se mantiene abierto las 24 horas del día, los 365 días del año, sólo queda fuera de servicio por causas de fuerza mayor u ocasiones muy especiales.

En 2018, tras la primera evaluación anual, Soo Yeon concluyó que la iniciativa BikePro había sido todo un éxito en la ciudad capital y podría replicarse en otras ciudades que lo requieran.

Soo Yeon es licenciada en derecho y ha sido ministro de movilidad los últimos cuatro años, especializándose en políticas públicas. Junto con su equipo, empezó la planeación para implementar BikePro en Busán, Incheon y Daegu. Establecieron los puntos importantes de cada ciudad donde sería útil una estación, ampliaron la capacidad de los servidores que sostienen el sistema de alquiler, incluso diseñaron una app para facilitar el registro y la renta de bicicletas desde los celulares.

Después de un mes de planeación, comenzaron a trabajar en la instalación de las estaciones, el señalamiento del carril designado para los ciclistas y la colocación de la señalética. Este proceso fluía sin problemas gracias a la experiencia que les había proporcionado la aplicación de este sistema en Seúl. No obstante, cuando llegaron al punto de la compra de bicicletas se dieron cuenta de que no podrían aplicar los mismos parámetros en todas las ciudades.

Para saber en cuántas bicicletas deben invertir, es necesario saber cuántas deben estar disponibles en cada estación. Así que convocaron a los técnicos que se encargan de administrar las unidades. En la reunión, todos comentaron que la distribución se ha hecho de manera empírica, aprendiendo con el tiempo cuáles son las horas pico y las horas muertas, cuáles son los mayores flujos, dónde suelen acumularse las bicicletas y dónde aumenta la demanda. Esto había funcionado en Seúl, pero para su expansión en otras ciudades necesitaban una predicción aproximada que les permitiera hacer una inversión de manera precisa.

Soo Yeon estuvo pensando un tiempo cómo obtener esta información, entonces recurrió a su agenda en busca del teléfono de Kwon Min, ministro de Análisis de Datos. Decidió esperar a la mañana siguiente para contactarla, así que dejó una nota con el número en medio de su escritorio.

En punto de las 10:00 am, tomó su teléfono y marcó la extensión de Kwon Min, quien atendió la llamada enseguida. Después de un breve saludo, quedaron en tener una junta por la tarde para que Soo Yeon pudiera plantearle la problemática y juntas buscar una solución.

Durante la junta, Soo Yeon explicó a detalle el plan de expandir BikePro, el estatus en el que se encontraban, así como, la necesidad de calcular la inversión que se requiere y administrar bien las unidades que se adquieran. Kwon Min, al oír las necesidades de Soo Yeon, pensó que podría generarse un sistema predictivo Machine Learning (ML) que les permita anticiparse a la demanda de bicicletas por día y hora, para que no se queden sin inventarió o al contrario, tener un excedente sin usar.

A Soo Yeon le pareció una buena propuesta y dejó en sus manos el desarrollo de este sistema. Kwon Min estuvo pensando durante un par de días a quién podía encargarle esta tarea, ella había dirigido anteriormente equipos de análisis para diferentes proyectos, así que tenía varias personas en mente, hasta que se decidió por Woo Young. Ella era una analista que trabajaba bajo su dirección desde hace 3 años, y que tenía dominio en estadística, aprendizaje de máquina e implementación de scripts en Python.

De camino a la institución, justo entrando en las instalaciones, Kwon Min se encontró con Woo Young, a quien le pidió que fuera a su oficina en cuanto pudiera. Woo Young llegó a su escritorio, terminó unos pendientes que tenía y se dirigió a la oficina de Kwon Min.

Kwon Min le contó sobre el proyecto BikePro, los resultados que había tenido en Seúl, y el plan de expansión que estaba en proceso, para el cual necesitaban que ella hiciera un prototipo de sistema de predicción de demanda de las bicicletas. Se esperaba que esta predicción se hiciera a nivel día y hora, considerando un horizonte de dos meses hacia el futuro,

pues es la granularidad que el equipo de Soo Yeon necesita para calcular la compra de unidades.

Woo Young tomó nota de la reunión y, mientras escuchaba los requerimientos, comenzó a pensar en la estrategia para resolver el problema. Le presentó a Kwon Min la propuesta de diseñar algoritmos y modelos de predicción que aborde series de tiempo básicos empleando ML, el cual es una rama de IA que dota a las máquinas la capacidad de aprender automáticamente, lo cual las hace muy eficaces para generar datos con base en predicciones. Esta idea le pareció muy buena a Kwon Min, así que aprobó el plan de Woo Young y le dio luz verde para trabajarlo.

Al avanzar en el desarrollo y revisar sus notas, Woo Young se dio cuenta de que necesitaba algunos datos sobre la demanda actual de bicicletas en Seúl, por lo que solicitó a su jefa los medios para obtener esta información. Ante la petición de Woo Young, Kwon Min agendó una cita con los técnicos de BikePro para el día siguiente a las 12:00 pm, para tener la información de primera mano y que no se les escapara ningún detalle.

Al día siguiente, en la reunión, los técnicos le proporcionaron a Woo Young todos los datos históricos que ellos habían obtenido con su experiencia, los cuales permitieron a Woo Young hacer una lista de los factores que pueden afectar la predicción. Entre los factores que más influían estaba el clima; pues si hay lluvias o frío, la gente prefiere otro tipo de transporte para resguardarse; los días festivos, reduce la cantidad de personas que se trasladan a sus centros de trabajo, ya que no todos optan por la opción de salir a pasear; y los días y horas en la semana, sobre este punto aún no habían llegado a un consenso, pues variaba dependiendo de la zona: cerca de escuelas y oficinas había picos muy notorios en los horarios de entrada y la salida, reduciendo a medio día,

pero en otras estaciones, cerca de parques o plazas, los fines de semana era cuando aumentaba exponencialmente el flujo de unidades.

Después de la reunión, Woo Young se dispuso a limpiar los datos obtenidos cuando le llegó un correo de Kwon Min, el cual le decía que querían que el prototipo de modelo con ML estuviera listo al terminar el sprint. Esto preocupó un poco a Woo Young, pues aparte de este proyecto tenía otras tareas que completar, pero confiaba en su experiencia trabajando con IA, así que continuó con su trabajo.

Además, en ese momento, el tiempo no era la principal preocupación de Woo Young. Al ver la información que disponía, se preguntaba ¿cómo podría identificar una afinidad entre las variables subjetivas y empíricas que le proporcionaron los técnicos para hacer proyecciones de demanda hacia el futuro? Esta duda se reflejaba en su rostro mientras comía, así que una colega le preguntó qué era lo que le tenía tan pensativa. Cuando Woo Young le contó, su colega le recomendó elaborar un análisis exploratorio de datos para detectar la relación entre los factores que afectan el fenómeno a predecir. Esta idea hizo sentido a Woo Young, quien dejó a medias el pollo frito que estaba comiendo y corrió a su oficina.

Empezó a desarrollar el análisis exploratorio de datos en Python, donde describió las variables y la forma en la que afectan al fenómeno por medio de sus distribuciones de probabilidad en términos de sus correlaciones. Para identificar la afinidad lineal entre dos variables empleó un análisis del coeficiente de correlación, probó con exploraciones univariadas, bivariadas y multivariadas. Una vez entendidas las relaciones, podría introducirlas como variables independientes o regresores.

Después de crear este análisis, planteó un modelo de aprendizaje de máquina, hizo correcciones tras ciertos errores derivados del sesgo y la

varianza, e incluyó frameworks para entrenar, probar y validar el modelo de aprendizaje.

Ya con el modelo perfeccionado, lo implementó en Python manipulando la información para atender los problemas predictivos, dividiendo los datos en conjuntos de prueba y entrenando el modelo con Sklearn para validación. Continuó con la creación, codificación y selección de variables a partir de los datos proporcionados por los técnicos. Utilizó estas variables para entrenar el modelo empleando Pipeline de Sklearn.

Para abordar las series de tiempo, usó APIs de Sklearn, probando con KNN, regresión lineal, árboles de decisión y bosques aleatorios. Por último, calibró los hiper parámetros y realizó una validación cruzada para seleccionar el modelo con mejor desempeño en Python.

Pasó un mes y Woo Young había terminado el prototipo. En la junta de cierre del sprint, Kwon Min solicitó a Woo Young un reporte de lo trabajado y la entrega del modelo que realizó. Woo Young preparó su entrega y, en menos de una hora estuvo en manos de Kwon Min.

Kwon Min puso a prueba el prototipo y encontró que todo funcionaba correctamente, así que programó la entrega con Soo Yeon, quien se mostró profundamente agradecida por el buen trabajo y la rapidez con la que se realizó.

El equipo de Soo Yeon empleó el modelo para hacer una predicción a dos meses con el fin de obtener un estimado de cuántas unidades se requieren por ciudad, sabiendo que, una vez que se implementara, podría presentar un margen de 200 unidades de error de raíz cuadrada media (RMSE).

Meses después, cuando se puso en marcha el proyecto BikePro, el equipo de administración utilizó un sistema de alertas tempranas que notifican la demanda futura de bicicletas. De manera que los técnicos puedan

optimizar el inventario: si se espera que la demanda sea alta en ciertas estaciones, se ajusta la distribución para asegurarse de tener suficientes unidades disponibles y así mejorar la rentabilidad, capacidad de crecimiento y, sobre todo, la experiencia del usuario al siempre encontrar bicicletas cuando las necesita.