

# MACHINE LEARNING CANVAS

## PREDICTION TASK



Tipo de tarea: Regresión.

Entidad predicha: Número de bicicletas alquiladas por hora (por ubicación).

Resultados posibles: Valores numéricos de demanda (entre 0 y ~800).

Momento de observación: Las salidas (número de rentas) se observan al cierre de cada hora en el sistema.

## DECISIONS



Las predicciones permiten planear redistribución diaria de bicicletas en función de zonas de alta y baja demanda.

En horarios pico, los operadores pueden asignar más unidades y personal en estaciones clave.

Las decisiones son asistidas, no automatizadas, pero se visualizan en paneles con alertas proactivas.

## VALUE PROPOSITION



Beneficiarios: Operadores del sistema de bicicletas compartidas en Seúl y usuarios del servicio.

Problemas abordados: Escasez o sobreabastecimiento de bicicletas en estaciones específicas; costos logísticos elevados; baja satisfacción del usuario.

Integración en el flujo de trabajo: Las predicciones se integran en el sistema de planificación logística diaria.

Interfaz de usuario: Dashboard interno para los operadores; reportes automáticos diarios/semanales.

## DATA COLLECTION



Inicial: Dataset histórico con sensores automáticos (CSV público).

Actualización: Diaria, desde sensores IoT en estaciones y APIs meteorológicas.

Estrategias: Automatización vía cron jobs y API calls; uso de pipelines de ingestión para mantener fresca sin costo excesivo.

## DATA SOURCES



Internos: Sistema de gestión de bicicletas (datos de renta por hora).

Externos:

- API meteorológico (OpenWeatherMap).
- Calendario oficial (días festivos y laborales).
- Sensores en estaciones para tráfico en tiempo real.

## IMPACT SIMULATION



Costo de error:

- Subestimación → Desabasto = usuarios sin servicio = pérdida de ingresos.
- Sobreestimación → Exceso = gastos logísticos innecesarios.

Simulación previa al despliegue: Pruebas históricas sobre fechas clave (festivos, clima extremo).

Criterios de despliegue: RMSE aceptable (< 50); mejora sobre baseline.

Restricciones de equidad: No aplicables directamente, pero se puede revisar equidad geográfica entre zonas.

## MAKING PREDICTIONS



Frecuencia: Batch diario, con posibilidad de actualizar predicciones cada hora.

Tiempo disponible: < 2 minutos por ciclo completo.

Recursos computacionales: Servidor cloud (EC2 pequeño o Lambda); cómputo ligero.

## Objetivo

Predecir la demanda de bicicletas compartidas en Seúl para mejorar la disponibilidad del servicio, optimizar la distribución de bicicletas y reducir costos logísticos. Esta solución beneficiará principalmente a los operadores del sistema y usuarios del servicio, y contribuirá a una movilidad urbana más sostenible.

## BUILDING MODELS



Cantidad de modelos: Uno principal (por estación o agrupamiento de zonas).

Actualización: Mensual o al detectar data drift.

Tiempo disponible: Hasta 10 minutos por reentrenamiento.

Recursos: CPU estándar, pipelines orquestados vía cron o Airflow.

## FEATURES



Representación:

- Hora del día
- Día de la semana
- Condiciones climáticas
- Visibilidad, lluvia, humedad
- Indicador de festivo

Transformaciones:

- One-hot encoding de categorías.
- Normalización de variables continuas.
- Combinaciones de clima + horario (ej. "lluvia lunes mañana").

## MONITORING



Métricas para usuarios y negocio:

- Desviación entre predicción y demanda real
- % de cumplimiento de demanda en estaciones clave

Frecuencia de revisión:

- Reducción de reclamos/logística innecesaria
- Semanal para métricas de negocio
- Diario para métricas operativas (error, desabasto)

