

MACHINE LEARNING CANVAS

PREDICTION TASK	DECISIONS	VALUE PROPOSITION	DATA COLLECTION	DATA SOURCES
<p>Tipo de tarea: Regresión.</p> <p>Entidad predicha: Número de bicicletas alquiladas por hora (por ubicación).</p> <p>Resultados posibles: Valores numéricos de demanda (entre 0 y ~800).</p> <p>Momento de observación: Las salidas (número de rentas) se observan al cierre de cada hora en el sistema.</p>	<p>Las predicciones permiten planear redistribución diaria de bicicletas en función de zonas de alta y baja demanda.</p> <p>En horarios pico, los operadores pueden asignar más unidades y personal en estaciones clave.</p> <p>Las decisiones son asistidas, no automatizadas, pero se visualizan en paneles con alertas proactivas.</p>	<p>Beneficiarios: Operadores del sistema de bicicletas compartidas en Seúl y usuarios del servicio.</p> <p>Problemas abordados: Escasez o sobreabastecimiento de bicicletas en estaciones específicas; costos logísticos elevados; baja satisfacción del usuario.</p> <p>Integración en el flujo de trabajo: Las predicciones se integran en el sistema de planificación logística diaria.</p> <p>Interfaz de usuario: Dashboard interno para los operadores; reportes automáticos diarios/semanales.</p>	<p>Inicial: Dataset histórico con sensores automáticos (CSV público).</p> <p>Actualización: Diaria, desde sensores IoT en estaciones y APIs meteorológicas.</p> <p>Estrategias: Automatización vía cron jobs y API calls; uso de pipelines de ingestión para mantener frescura sin costo excesivo.</p>	<p>Internos: Sistema de gestión de bicicletas (datos de renta por hora).</p> <p>Externos:</p> <ul style="list-style-type: none"> ● API meteorológico (OpenWeatherMap). ● Calendario oficial (días festivos y laborales). ● Sensores en estaciones para tráfico en tiempo real.
IMPACT SIMULATION	MAKING PREDICTIONS	BUILDING MODELS	FEATURES	
<p>Costo de error:</p> <ul style="list-style-type: none"> ● Subestimación → Desabasto = usuarios sin servicio = pérdida de ingresos. ● Sobreestimación → Exceso = gastos logísticos innecesarios. <p>Simulación previa al despliegue: Pruebas históricas sobre fechas clave (festivos, clima extremo).</p> <p>Criterios de despliegue: RMSE aceptable (< 50); mejora sobre baseline.</p> <p>Restricciones de equidad: No aplicables directamente, pero se puede revisar equidad geográfica entre zonas.</p>	<p>Frecuencia: Batch diario, con posibilidad de actualizar predicciones cada hora.</p> <p>Tiempo disponible: < 2 minutos por ciclo completo.</p> <p>Recursos computacionales: Servidor cloud (EC2 pequeño o Lambda); cómputo ligero.</p>	<p>Objetivo</p> <p>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.</p>	<p>Cantidad de modelos: Uno principal (por estación o agrupamiento de zonas).</p> <p>Actualización: Mensual o al detectar data drift.</p> <p>Tiempo disponible: Hasta 10 minutos por reentrenamiento.</p> <p>Recursos: CPU estándar, pipelines orquestados vía cron o Airflow.</p>	<p>Representación:</p> <ul style="list-style-type: none"> ● Hora del día ● Día de la semana ● Condiciones climáticas ● Visibilidad, lluvia, humedad ● Indicador de festivo <p>Transformaciones:</p> <ul style="list-style-type: none"> ● One-hot encoding de categorías. ● Normalización de variables continuas. ● Combinaciones de clima + horario (ej. "lluvia lunes mañana").
MONITORING				
	<p>Métricas para usuarios y negocio:</p> <ul style="list-style-type: none"> ● Desviación entre predicción y demanda real ● % de cumplimiento de demanda en estaciones clave 	<ul style="list-style-type: none"> ● Reducción de reclamos/logística innecesaria <p>Frecuencia de revisión:</p> <ul style="list-style-type: none"> ● Semanal para métricas de negocio ● Diario para métricas operativas (error, desabasto) 		

