











# MACHINE LEARNING CANVAS

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