**Universidad ORT Uruguay**

Facultad de Ingeniería

**Machine learning en producción**

**Obligatorio**

**Autores:**

Gonzalo de León - 158545

Martin Ljubicic - 154630

Joaquín Oldan - 159579

**Docentes:**

Federico Zaiter

Matías Sorozabal

**2023**

Tabla de contenido

[I. Introducción 4](#_Toc136772591)

[A. Contexto del estudio 4](#_Toc136772592)

[B. Objetivos 4](#_Toc136772593)

[C. Metodología de investigación 4](#_Toc136772594)

[II. Fundamentos de Machine Learning 5](#_Toc136772595)

[A. Definición de Machine Learning 5](#_Toc136772596)

[B. Tipos de algoritmos de Machine Learning 5](#_Toc136772597)

[C. Importancia del Machine Learning en la producción 5](#_Toc136772598)

[III. Preparación de los datos 6](#_Toc136772599)

[A. Recopilación de datos 6](#_Toc136772600)

[B. Limpieza y preprocesamiento de datos 6](#_Toc136772601)

[C. Selección de características 6](#_Toc136772602)

[IV. Modelado y entrenamiento 7](#_Toc136772603)

[A. Selección de algoritmos de Machine Learning 7](#_Toc136772604)

[B. División de conjuntos de datos 7](#_Toc136772605)

[C. Entrenamiento del modelo 7](#_Toc136772606)

[V. Implementación en producción 8](#_Toc136772607)

[A. Consideraciones de infraestructura 8](#_Toc136772608)

[B. Integración del modelo en un sistema de producción 8](#_Toc136772609)

[C. Monitoreo y mantenimiento del modelo en producción 8](#_Toc136772610)

[VI. Casos de estudio 9](#_Toc136772611)

[A. Descripción del caso de estudio 1 9](#_Toc136772612)

[B. Resultados y lecciones aprendidas del caso de estudio 1 9](#_Toc136772613)

[C. Descripción del caso de estudio 2 9](#_Toc136772614)

[D. Resultados y lecciones aprendidas del caso de estudio 2 9](#_Toc136772615)

[VII. Desafíos y recomendaciones 10](#_Toc136772616)

[A. Desafíos comunes en la implementación de Machine Learning en producción 10](#_Toc136772617)

[B. Recomendaciones para superar los desafíos 10](#_Toc136772618)

[VIII. Conclusiones 11](#_Toc136772619)

[A. Resumen de los hallazgos 11](#_Toc136772620)

[B. Contribuciones del estudio 11](#_Toc136772621)

[C. Áreas para futuras investigaciones 11](#_Toc136772622)

# I. Introducción

## A. Contexto del estudio

El Machine Learning surge como una necesidad de estudiar los datos obtenidos de cualquier ámbito para obtener información relevante para problemas planteados o para mejorar sistemas. La toma de decisiones automatizada, la automatización de procesos aporta mucho valor agregado en diferentes ámbitos.

Como se pudo observar en las presentaciones realizadas por el equipo de diversos papers, existen cada vez mas casos de uso exitosos en la materia, en la presentación de los integrantes del grupo sobre Michelangelo, se expuso que no solo es importante los resultados directos del aprendizaje, sino la metodología aplicada para llegar a ellos. Esta herramienta proporciona buenos resultados y ayudo a organizar los grupos de trabajo.

La automatización es vista como como una oportunidad de mejora para las organizaciones y permitirá redirigir sus esfuerzo a trabajos de mejora constante de estas metodologías, logrando mejorar sus procesos.

En este trabajo se buscará incursionar en estas metodologías para la implementación de un producto que pueda ser llevado a producción haciendo incapie mas en el E2E que en sus diferentes partes.

## B. Objetivos

El objetivo general del proyecto consiste en lograr un E2E de un sistema que obtenga datos de una pagina web de venta de propiedades y que sea capaz de aprender a predecir un parámetro de estas relevante, en este caso se apunta a poder distinguir la categoría del inmueble según su valor.

En principio esta variable esta directamente relacionada con el precio, aunque se podria llegar a paralelizar una categoría de valor en función de las necesidades de quien busca adquirir la vivienda. Esto ya es un desafío mas ambicioso y no será el foco del presente trabajo.

Se logran identificar objetivos mas específicos, apuntando a etapas de proceso general.

Obtener un modelo que logre analizar imágenes como datos tabulares de los elementos de estudio, si bien no hay un enfoque sobre obtener los mejores resultados, si se buscara un mínimo desempeño en esta etapa del modelo, que sirva para obtener resultados bastante aproximados del problema.

Análisis de datos, comprensión de cuales son mas relevantes y cuales no, cuales aportan mas al target que se busca aprender para luego inferir con buena precisión en datos de testeo.

Desempeño del sistema, en velocidad y escalabilidad, un buen sistema debe ser capaz de proporcionar una buena respuesta ante múltiples consultas.

Evaluar el impacto de la solución completa en cuanto a su utilidad real y la aplicabilidad a un caso real de estudio, si bien se puede considerar un objetivo secundario, seria bueno tener una referencia a su utilidad, por su valor agregado a algún proceso determinado para un caso de estudio puntual o general.

Comprender que mejoras se pueden proporcionar a futuro en la aplicación para apuntar a un mercado de clientes reales y su posibilidad de venta.

## C. Metodología de investigación

# II. Fundamentos de Machine Learning

## A. Definición de Machine Learning

## B. Tipos de algoritmos de Machine Learning

## C. Importancia del Machine Learning en la producción

# III. Preparación de los datos

## A. Recopilación de datos

## B. Limpieza y preprocesamiento de datos

## C. Selección de características

# IV. Modelado y entrenamiento

## A. Selección de algoritmos de Machine Learning

## B. División de conjuntos de datos

## C. Entrenamiento del modelo

# V. Implementación en producción

## A. Consideraciones de infraestructura

## B. Integración del modelo en un sistema de producción

## C. Monitoreo y mantenimiento del modelo en producción

# VI. Casos de estudio

## A. Descripción del caso de estudio 1

## B. Resultados y lecciones aprendidas del caso de estudio 1

## C. Descripción del caso de estudio 2

## D. Resultados y lecciones aprendidas del caso de estudio 2

# VII. Desafíos y recomendaciones

## A. Desafíos comunes en la implementación de Machine Learning en producción

## B. Recomendaciones para superar los desafíos

# VIII. Conclusiones

## A. Resumen de los hallazgos

## B. Contribuciones del estudio

## C. Áreas para futuras investigaciones

FIN DEL DOCUMENTO