Proyecto de Inteligencia Artificial Bootcamp IAp191

Ejecutor: V´ıctor Aguilar

17 de septiembre de 2025

´Indice

Estructura de Carpetas del Proyecto 3 1. Introducci´on 4 2. Definici´on del Problema 4 3. Revisi´on de Literatura 4

4. Metodolog´ıa 5 4.1. Datos . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5 4.2. Modelos . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5 4.3. Evaluaci´on . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

5. Resultados 5 6. Discusi´on 6 7. Conclusiones 6

1

Bootcamp IAp191 Ejecutor: V´ıctor Aguilar 8. Referencias 6

2

Bootcamp IAp191 Ejecutor: V´ıctor Aguilar Estructura de Carpetas del Proyecto

Para mantener orden y reproducibilidad, se recomienda la siguiente estructura:

project\_name/

README.md # Explica prop´osito, uso y pasos de ejecuci´on requirements.txt # Librer´ıas necesarias (p.ej. scikit-learn, keras)

data/

raw/ # Datos originales (no se modifican nunca) processed/ # Datos limpios listos para modelos

notebooks/ # Jupyter para an´alisis exploratorio y prototipos results/

metrics/ # Reportes en CSV/JSON con accuracy, F1, etc. figures/ # Gr´aficas exportadas (curvas, matrices de confusi´on) docs/ # Documento final en LaTeX/PDF

Nota para estudiantes: - Nunca sobrescribir los datos en raw/. - Documentar cambios de preprocesamiento en processed/. - Guardar gr´aficos y m´etricas en results/. - En notebooks/ mantener experimentos claros con celdas ordenadas.

3

Bootcamp IAp191 Ejecutor: V´ıctor Aguilar 1 Introducci´on

Aqu´ı el estudiante explica el problema en un lenguaje sencillo:

¿Qu´e tema eligi´o?

¿Por qu´e le interesa?

¿A qui´en beneficia resolverlo?

Ejemplo: “El an´alisis de sentimientos en rese˜nas de pel´ıculas es ´util para que plataformas de cine automaticen la clasificaci´on de opiniones y entiendan la experiencia del usuario. Este proyecto explora un modelo de aprendizaje autom´atico para predecir si una rese˜na es positiva o negativa”.

2 Definici´on del Problema

Se formula el problema de forma clara y medible.

Objetivo general.

2–3 objetivos espec´ıficos.

Ejemplo: - Objetivo general: Desarrollar un clasificador de rese˜nas de pel´ıculas en positivas o negativas. - Objetivos espec´ıficos: 1. Preparar un conjunto de datos de rese˜nas en espa˜nol e ingl´es. 2. Entrenar un modelo cl´asico (Naive Bayes). 3. Entrenar un modelo profundo (LSTM) y comparar resultados.

3 Revisi´on de Literatura

Los estudiantes deben consultar al menos 2 fuentes (libros, art´ıculos, manuales). Ejemplo de redacci´on sin cita: “Diversos estudios han demostrado que los modelos de redes neu ronales recurrentes capturan mejor las dependencias en texto que los algoritmos basados en conteo de palabras. Por otro lado, los modelos probabil´ısticos simples siguen siendo competitivos cuando los datos son limitados”.

4

Bootcamp IAp191 Ejecutor: V´ıctor Aguilar

**4 Metodolog´ıa**

4.1 Datos

Indicar origen del dataset.

Explicar c´omo se dividieron (train/test).

Describir pasos de limpieza.

Ejemplo: “El dataset IMDB contiene 50 mil rese˜nas, divididas en 25 mil para entrena miento y 25 mil para prueba. Se realiz´o tokenizaci´on y truncamiento a 200 palabras por rese˜na”.

**4.2 Modelos**

Los estudiantes deben entrenar: - Un modelo cl´asico de ML. - Un modelo de Deep Lear ning.

Ejemplo: “Se entren´o un Naive Bayes con suavizado de Laplace y un modelo LSTM con embedding de 128 dimensiones y 64 unidades recurrentes”.

**4.3 Evaluaci´on**

Explicar m´etricas y c´omo se usaron.

Ejemplo de tabla:

Modelo Accuracy Precisi´on Recall F1

Naive Bayes 0.85 0.84 0.85 0.85

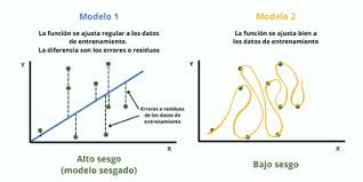
LSTM 0.90 0.89 0.91 0.90

**5 Resultados**

Se muestran gr´aficas y m´etricas finales.

Ejemplo: “El modelo LSTM super´o al Naive Bayes en todas las m´etricas, especialmente en F1”.

5

Bootcamp IAp191 Ejecutor: V´ıctor Aguilar Curva de aprendizaje del LSTM.

6 Discusi´on

Aqu´ı los estudiantes analizan cr´ıticamente:

¿Por qu´e un modelo funcion´o mejor?

¿Qu´e limitaciones tuvo el experimento?

¿Qu´e mejoras proponen?

Ejemplo: “El LSTM obtuvo mejor rendimiento porque captura dependencias entre pa labras. Sin embargo, requiere m´as tiempo de entrenamiento y recursos computacionales”.

7 Conclusiones

Se resumen aprendizajes en 4–6 l´ıneas.

Ejemplo: “Se logr´o clasificar rese˜nas con un nivel aceptable de precisi´on. Los algoritmos de Deep Learning muestran mejores resultados que los cl´asicos, aunque requieren mayor infraestructura. Una futura extensi´on es probar transformers”.

**8 Referencias**

Aqu´ı los estudiantes deben redactar sus fuentes seg´un el formato que se indique (APA 7 u otro).

6