



Universidad de Santiago de Chile

Facultad de Ingeniería

Departamento de Ingeniería Informática

Proyecto de Optimización I-2025

FECHA: 24-04-2024

NOMBRE DE PROFESOR: Víctor Parada

NOMBRE DEL AYUDANTE: Felipe Gatica, Matías Yáñez

Objetivo: Construir un modelo matemático de inteligencia artificial que se ajuste a un dataset dado. Parte 1: Caracterización del problema

Cada grupo de estudiantes recibe un conjunto de datos de estudio seleccionados de kaggle.com. Se debe obtener el conjunto parámetros de un modelo de red neuronal multicapa que se ajuste al conjunto de datos dados. La etapa de solución del problema de optimización correspondiente (entrenamiento) se debe llevar a cabo utilizando los siguientes algoritmos de optimización: SGD, RMSprop, Adam, Lamb. Tales algoritmos están disponibles para ser utilizados con la biblioteca Keras.

Parte 2: Implementar funciones de penalidad

1.1 Selección de funciones de penalidad:

- a) Implemente dos funciones de penalidad que eviten el sobreentrenamiento y realice un análisis comparativo.
- b) Medir métricas en el conjunto de Test, para distintas funciones de penalidad y distintos algoritmos de optimización.

1.2 Codificación en Python utilizando Keras:

- a) El código debe estar bien estructurado y documentado, facilitando la comprensión del proceso de resolución y garantizando la replicabilidad del experimento por terceros. b) Se debe asegurar que el código sea capaz de resolver la instancia de manera eficiente, produciendo resultados correctos en un tiempo razonable.

1.3 Resultados Numéricos:

- a) Presentar los resultados numéricos obtenidos al entrenar el modelo con el dataset asignado.
- b) Incluir en los resultados métricas como:
 - i. El valor de la función objetivo.

- ii. Tiempo computacional requerido para el entrenamiento.

Los resultados deben ser presentados en un documento que obedece al formato paper en Latex de IEEE que se encuentra en el siguiente link.

<https://www.ieee.org/conferences/publishing/templates.html>

1.4 Informe en Formato IEEE:

- a) El informe debe ser presentado en un artículo de máximo 4 páginas utilizando el formato estándar IEEE.
- b) Debe incluir las siguientes secciones:
 - iii. Título
 - iv. Autores
 - v. Resumen: Breve descripción del problema y los resultados obtenidos.
 - vi. Introducción: Explicación del contexto del problema, relevancia del mismo y objetivos del proyecto.
 - vii. Revisión de la Literatura: Análisis de trabajos previos relacionados con la optimización en el sector salud.
 - viii. Modelos Matemáticos: Descripción de la formulación matemática del problema y las instancias utilizadas.
 - ix. Resultados: Presentación y análisis de los resultados obtenidos, incluyendo las métricas computacionales.
 - x. Conclusión: Reflexiones finales sobre el trabajo realizado y posibles mejoras.
 - xi. Bibliografía: Lista de las fuentes y referencias utilizadas en el proyecto.

Criterios de Evaluación:

1. Comprensión del dataset asignado.
2. Correcta formulación matemática del problema.
3. Calidad y diversidad de los ejemplos presentados (funciones de penalidad).
4. Funcionamiento y eficiencia del código en Python.
5. Claridad y profundidad del análisis de los resultados.
6. Calidad del artículo en formato IEEE, incluyendo la redacción y organización del contenido.
7. Presentación: Realizar una presentación de 20 minutos que describa el problema seleccionado considerando los siguientes aspectos:
 - Descripción del problema a partir del dataset asignado.
 - Descripción de la forma en que este problema se resuelve en la actualidad. -
 - Estimación del impacto cuantitativo de la solución.
 - Estimación del beneficio potencial en el contexto del dataset asignado
 - Identificación de la toma de decisiones en el ámbito del dataset asignado. -
 - Identificar variables y parámetros considerados en el dataset asignado.
 - Identificar criterios de optimización involucrados en el dataset asignado. -

Identificar las principales limitaciones que condicionan el problema.

- Formulación matemática del problema de optimización.

- Al final de la presentación se debe hacer entrega del documento usado en la presentación.

Fechas de Entrega:

Presentación Final: 4 y 5 - 06 - 2025

Equipo de trabajo: El proyecto debe ser realizado en grupos de no más de dos personas.

Parte 3: Asignación de grupos

Grupo 1: Iris Species, <https://www.kaggle.com/datasets/uciml/iris>

Grupo 2: Breast Cancer Wisconsin,

<https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data>

Grupo 3: Pima Indians Diabetes Database,

<https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database>

Grupo 4: Titanic Dataset, <https://www.kaggle.com/datasets/yasserh/titanic-dataset>

Grupo 5: Default of Credit Card, <https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset>

Grupo 6: Bank Marketing Dataset,

<https://www.kaggle.com/datasets/hariharanpavan/bank-marketing-dataset-analysis-classification>

Grupo 7: Heart Disease Dataset UCI,

<https://www.kaggle.com/datasets/ketangangal/heart-disease-dataset-uci>

Grupo 8: Mushroom Classification, <https://www.kaggle.com/datasets/uciml/mushroom-classification>

Grupo 9: Telco Customer Churn, <https://www.kaggle.com/datasets/blatchar/telco-customer-churn>

Identificación de grupos

[Redacted group identification information]

