

ISWZ3402 **Inteligencia Artificial II**

Ejercicio práctico

Elaborado por: **Mario González**

RESULTADO DE APRENDIZAJE DE LA CARRERA:

RC7. Adquiere y aplica nuevos conocimientos según sea necesario, utilizando estrategias de aprendizaje apropiadas.

Indicadores de desempeño:

- Indaga técnicas y herramientas necesarias acorde al problema planteado
- Argumenta las técnicas y herramientas disponibles
- Aplica las técnicas y herramientas pertinentes al problema planteado

OBJETIVO PROPUESTO DE LA CONSIGNA:

The objective of this homework is to develop and implement a quadratic regression model to better understand and predict the relationship between variables where linear models may fall short. Through this assignment, students will gain hands-on experience in data preprocessing, model building, and evaluation within the context of quadratic regression. This includes applying the theoretical concepts of polynomial regression, enhancing their proficiency in Python programming, and leveraging libraries such as NumPy and scikit-learn. By completing this task, students will be able to analyze complex datasets, interpret the results of quadratic regression, and compare its performance with linear regression models, thereby deepening their understanding of machine learning techniques and their applications.

INDICACIONES:

Implement a polynomial (quadratic) regression

Model 1: Analytical Solution (Quadratic Regression)

For this model, you'll use libraries like sklearn, scipy, or statsmodel to perform quadratic regression. You can replicate the approach from linear regression, but instead of fitting a linear function, you'll fit a quadratic function. Here's a general outline:

1. **Import necessary libraries:** Import sklearn, scipy, or statsmodel for quadratic regression.
2. **Load the dataset:** Use the provided dataset:

https://raw.githubusercontent.com/Machine1314/natural_computing/main/data_regression.csv
3. **Prepare the data:** Preprocess the data as needed (scaling, splitting into training and testing sets).
4. **Fit the quadratic regression model:** Use the chosen library to fit the quadratic regression model to the training data.
5. **Evaluate the model:** Assess the model's performance using appropriate metrics.
6. **Compare with other models:** Keep the results for comparison with Models 2 and 3.
7. **Check:** https://marsgr6.github.io/presentations/curve_fitting_and_interpolation.slides.html#/4

Model 2: Numerical Solution (Gradient Descent)

This model involves finding the coefficients a_0 , a_1 , and a_2 for the quadratic regression using gradient descent. Here's what you'll do:

1. **Define the cost function:** Define a cost function (e.g., mean squared error) for quadratic regression.
2. **Implement gradient descent:** Write functions to perform gradient descent to minimize the cost function.
3. **Scale the data:** Scale the input features if needed.
4. **Fit the model using gradient descent:** Use the implemented gradient descent algorithm to find the optimal coefficients.
5. **Evaluate the model:** Assess the model's performance using appropriate metrics.
6. **Compare with Model 1:** Compare the results obtained from gradient descent with those from the analytical solution.
7. **Check:** https://deepnote.com/workspace/mario-gonzalez-911d-e512259f-42a9-4514-8972-e208f19e4b48/project/ai2202210-cce3455d-da08-4c10-a6fe-39f2a30c6a51/notebook/06_gradient_descent_202420-4b60a63aaa874332a84a86908d9de75d

Model 3: MLP (Multi-layer Perceptron)

In this model, you'll use a multi-layer perceptron (MLP) for quadratic regression. Here's what you'll do:

1. **Import necessary libraries:** Import libraries for building and training MLPs.

2. **Load and preprocess the data:** Use the provided dataset or the one from the dataset above. Scale the data if necessary.
3. **Build the MLP model:** Define a neural network architecture suitable for quadratic regression.
4. **Find best parameters:** Use techniques like grid search or random search to find the best hyperparameters for the MLP.
5. **Train the model:** Train the MLP model on the training data.
6. **Check for overfitting:** Monitor the training and validation loss to identify overfitting.
7. **Evaluate the model:** Assess the model's performance using appropriate metrics.
8. **Compare with other models:** Keep the results for comparison with Models 1 and 2.
9. **Check:** https://nbviewer.org/github/marsgr6/ml-online/blob/main/ann_dnn.ipynb

After implementing each model, you can compare their performance using metrics like mean squared error, R-squared, etc. To get back to the original scale for comparison, you'll need to reverse any scaling transformations applied to the predictions.

General instructions

- (-1 pts) Specify the group members in the first cell of the notebook.
 - ☐ Give an itemized list: each member in a line.
- (-1 pts) Deliver two files into a zip folder with the following name format:
 - ☐ P2_Project_IA202410_II1II2II3.zip: with two files: html (**or pdf**) and ipynb.
 - ☐ For example: Nikita Martínez, Nestor Lozada, Damián Briones, results in a file name: P2_Project_IA202410_NM1NL2DB3.zip.
- (-1 pts) **All group members should upload the same file**
- The report (notebook) will be evaluated as:
 - ☐ (-5 pts) Poor: only the code was delivered.
 - ☐ (-5 pts) Fair: commented code, with sections.
 - ☐ (-2 pts) Good: commented code, with sections and explanations of each part of the process.
 - ☐ (-0 pts) Excellent: commented code, with sections and explanations of each part of the process including formulas, research, and explanation of your ML model, the dataset exploration, the problem you want to solve, etc. Figures with labels, titles legends, etc.

FORMA DE TRABAJO:

La propuesta se la desarrollará en grupos de máximo de 5 integrantes.

ESPECIFICACIONES DE ENTREGA:

El estudiante debe entregar un informe completo y detallado en formato ipynb (jupyter notebook) donde se detalle cada una de las fases. Debe exportar el informe a html o pdf, y adjuntar junto con el notebook en ipynb.

Identifica necesidades de aprendizaje

- Define el problema de muestreo: random vs. Kmeans.
- Represente de forma gráfica los resultados comparando ambas formas de muestreo

Selecciona fuentes de información

- Investiga referencias que describen el uso de los modelos de ML seleccionados

Aplica las técnicas y herramientas pertinentes al problema planteado

- Realiza el preprocesamiento de los datos de entrada (si fuese necesario)
- Implementa el escalado adecuado de los datos (recuerde que algoritmo de ML tiene su particularidad de procesar el input)
- Implementa los modelos de ML para el problema y realiza los experimentos indicados
- Describe la solución (combinación de hiperparámetros) con su respectiva justificación
- Conclusiones

RÚBRICA:

CRITERIOS	EXCELENTE	MUY BUENO	BUENO	REGULAR	INSUFICIENTE
IDENTIFICA NECESIDADES DE APRENDIZAJE	Identifica necesidades de aprendizaje de manera autónoma y aplica estrategias apropiadas y relevantes que le permiten ampliar su conocimiento.	Identifica necesidades de aprendizaje de manera autónoma y aplica estrategias apropiadas que le permiten ampliar su conocimiento.	Identifica necesidades de aprendizaje de manera autónoma y aplica estrategias generales que le permiten ampliar su conocimiento.	Requiere apoyo para identificar necesidades de aprendizaje y aplica estrategias poco apropiadas que le permiten ampliar su conocimiento.	Requiere apoyo para identificar necesidades de aprendizaje y no aplica estrategias que le permiten ampliar su conocimiento.
SELECCIONA FUENTES DE INFORMACIÓN	Selecciona de manera proactiva fuentes de información adicional y persigue permanentemente experiencias educativas más allá de los requerimientos de su entorno de aprendizaje.	Selecciona de manera activa fuentes de información adicional y persigue permanentemente experiencias educativas más allá de los requerimientos de su entorno de aprendizaje.	Selecciona de manera básica, aunque apropiada , fuentes de información adicional y persigue regularmente experiencias educativas más allá de los requerimientos de su entorno de aprendizaje.	Selecciona de manera básica y poco apropiada fuentes de información adicional y persigue ocasionalmente experiencias educativas más allá de los requerimientos de su entorno de aprendizaje.	No selecciona fuentes de información adicional y no persigue experiencias educativas más allá de los requerimientos de su entorno de aprendizaje.
APLICA DESTREZAS Y CONOCIMIENTOS	Aplica de una manera innovadora (nueva y creativa) las destrezas y conocimientos, demostrando comprensión y excelente	Aplica de una manera adecuada las destrezas y conocimientos, demostrando comprensión y un óptimo desempeño	Aplica de una manera básica las destrezas y conocimientos, demostrando comprensión y buen desempeño frente a nuevas situaciones.	Aplica de una manera parcial las destrezas y conocimientos, demostrando poca comprensión y un desempeño regular	Aplica de forma errónea las destrezas y conocimientos, demostrando un desempeño deficiente frente a nuevas situaciones.

FACULTAD DE INGENIERÍA Y CIENCIAS APLICADAS

	desempeño frente a nuevas situaciones.	frente a nuevas situaciones.		frente a nuevas situaciones.	
--	--	---------------------------------	--	---------------------------------	--