Ramos, Kristel N.

BSCS-3B

Machine Problem No. 1

Fundamentals of Machine Learning

Mini-task

What is the input (features)?

- The input features are the numbers that describe the flower, like sepal length, sepal width, petal length, and petal width. These are the measurements we use to predict the type of flower.

What is the output (label)?

- The output is the kind of iris flower, which can be Setosa, Versicolor, or Virginica. It's what the model tries to guess from the features.

Is this supervised or unsupervised learning?

- This is supervised learning because the dataset already has answers (labels) for the flowers. The model learns by matching the features with the correct species.

Evaluation & Reflection

Evaluate model with different metrics (Classification).

- Using a confusion matrix, we can see how many predictions were correct or wrong. Precision and recall help us know how well the model identifies the correct flower types.

Evaluate model with different metrics (Regression).

- For regression, we use RMSE (Root Mean Squared Error) which shows how close the predictions are to the actual values. A smaller RMSE means the model is more accurate.

Discuss ML challenges: overfitting, underfitting, and bad data.

- Overfitting happens when the model memorizes the training data too much and fails on new data. Underfitting is when the model is too simple and misses important patterns. Bad data like missing or wrong values can confuse the model and make predictions worse.

What would happen if the dataset had missing or wrong values?

- If the dataset had missing or wrong values, the model would have trouble learning correctly. This could lead to low accuracy and unreliable predictions.

How does this relate to real-world ML applications?

- In real life, data is not always clean, like errors in medical records or forms. Machine learning needs proper data cleaning to give useful and trustworthy results.

Short reflection (3–5 sentences): o What ML type did you use? o What challenge might affect the model?

- I used supervised learning with the Iris dataset because the model was trained using features and their correct labels. The type of ML applied was classification, since the goal was to predict the flower species. A challenge that might affect the model is overfitting, where the model memorizes the training data and struggles with new data. Another challenge is having bad or missing data, which can make the predictions less accurate.