Fundamentos de Aprendizagem Automática 25/26

Universidade de Aveiro

Brain

TUMOR CLASSIFICATION

Carolina Silva, 113475, MEI

Matilde Teixeira, 108193, MECT

Motivation

01

HIGH mortality:

most dangerous types of cancer when not detected early.

02

Need for Better Diagnostics

ML provides fast, consistent and accurate tumor classification.

03

Importance of MRI:

MRI offers detailed brain imaging (safest, non-invasive method for detecting abnormalities and early-stage tumors).

04

Goal:

Reliable binary classifier capable of distinguishing tumor vs. no-tumor MRI images with high accuracy.

Brain Tumor MRI Dataset (Kaggle)

- Original classes: glioma, meningioma, pituitary, no tumor
    - Merged into: tumor vs. no tumor
    - Total images: 6726
        - Training: 5521
        - Testing: 1205
- Final resolution: 128×128
- Moderately imbalanced dataset

State of the Art

S: Classical ML: SVM, LR, handcrafted features

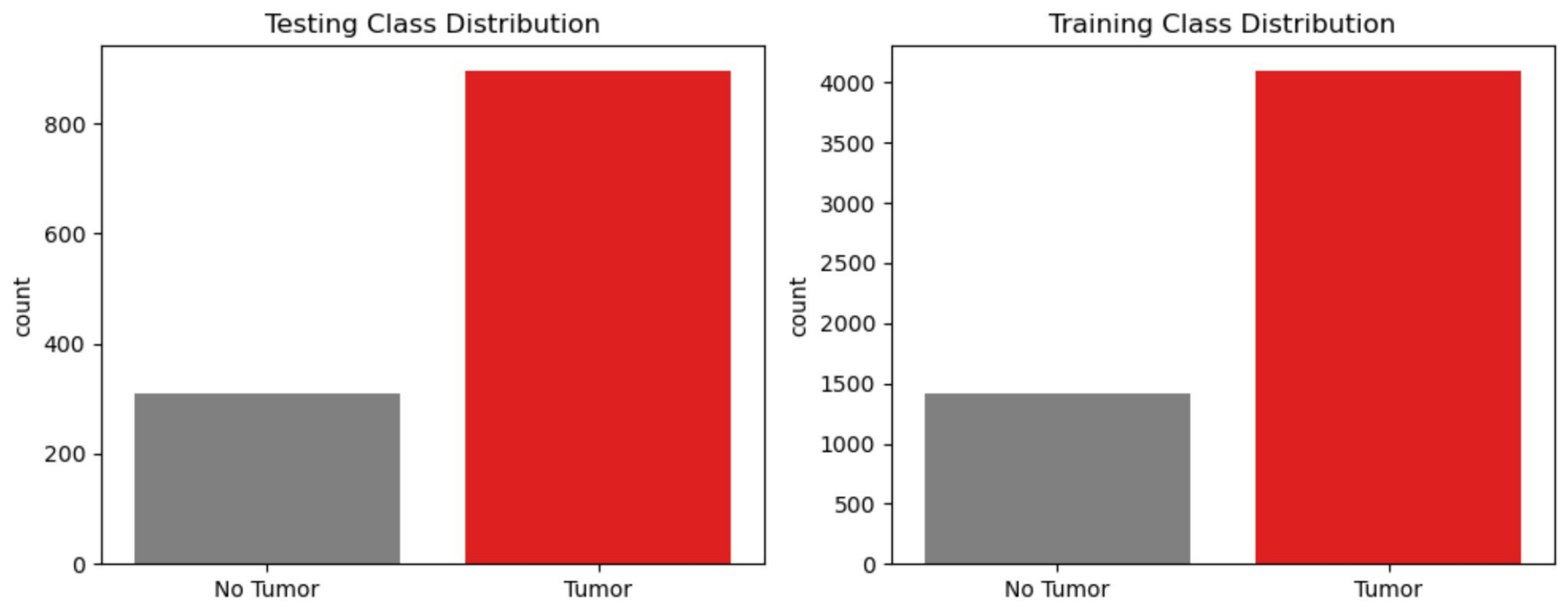
O: Deep learning: CNNs achieve >98–99%

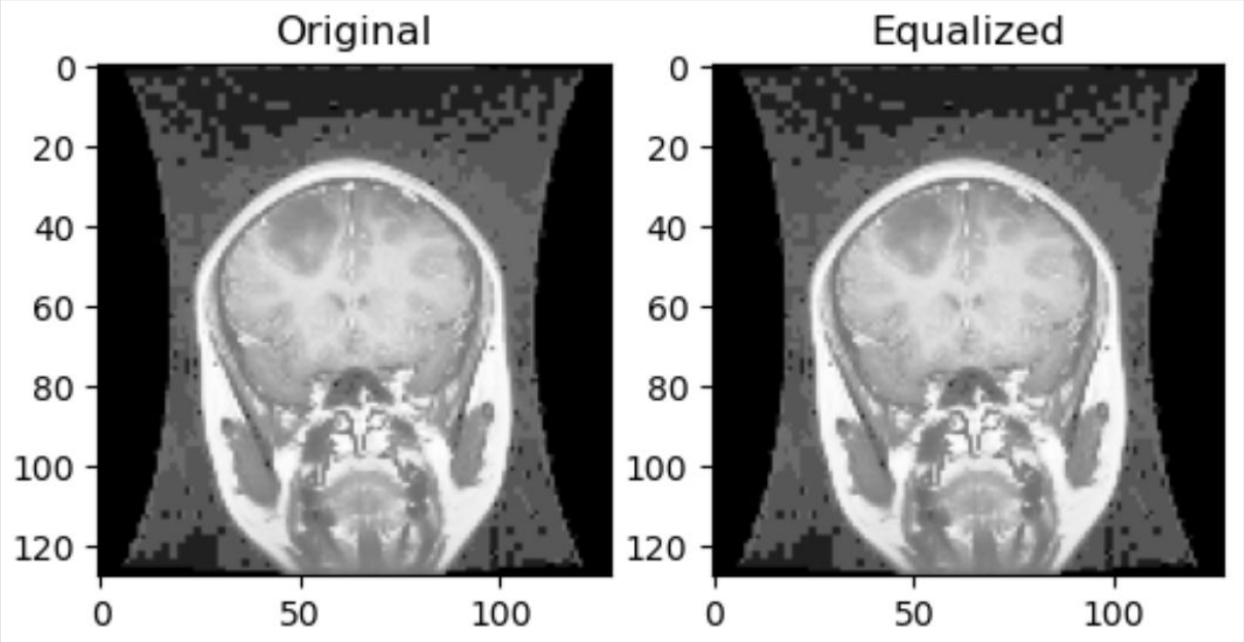
T: Hybrid CNN-SVM models widely used.

A: Challenges: imbalance, interpretability, variability

Pre-processing Pipeline

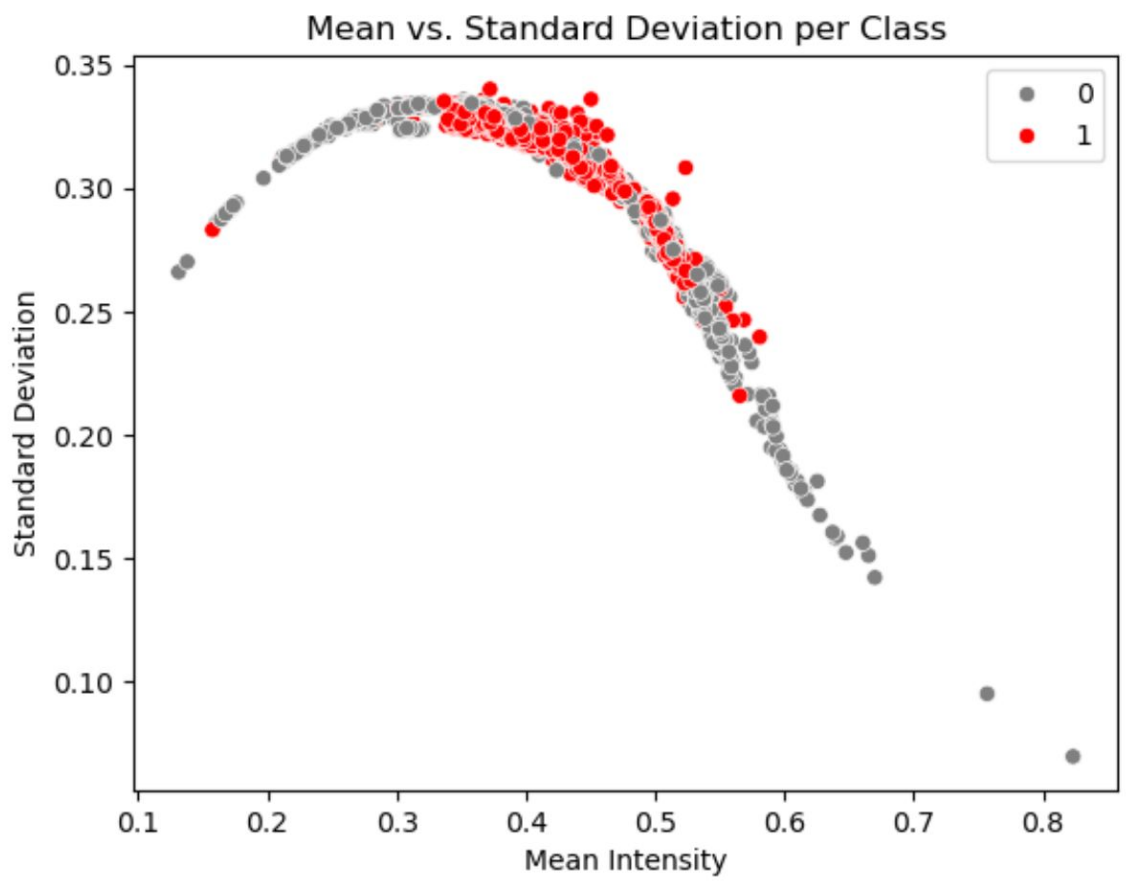
- Duplicate removal (297 images).
- Grayscale conversion.
- Resizing to 128×128.
- Histogram equalization for contrast improvement.
- Min–max normalization to [0,1].



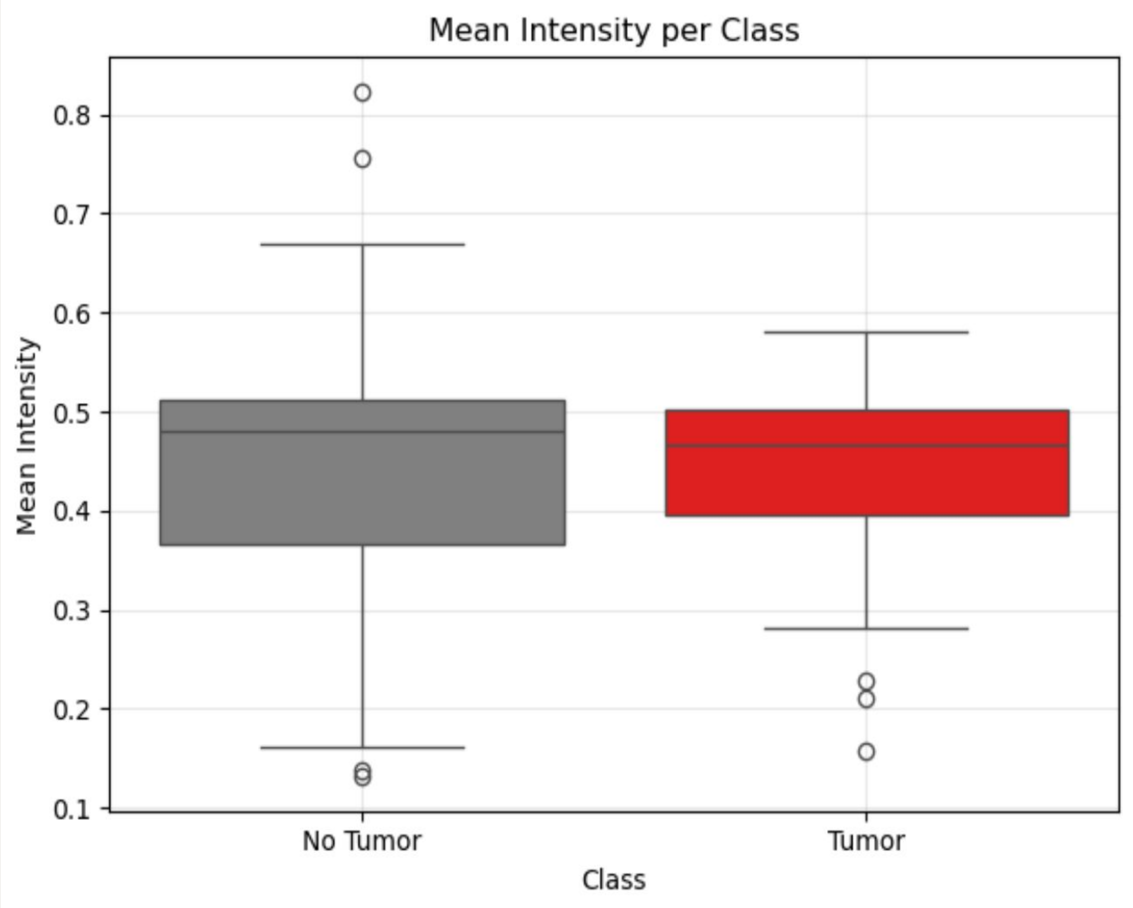


MEAN VS STANDARD DEVIATION

Summary: dataset is consistent and well balanced after processing.



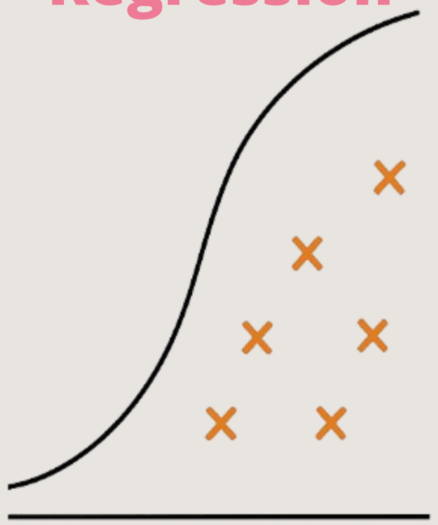
Mean & standard deviation comparison (train/test)



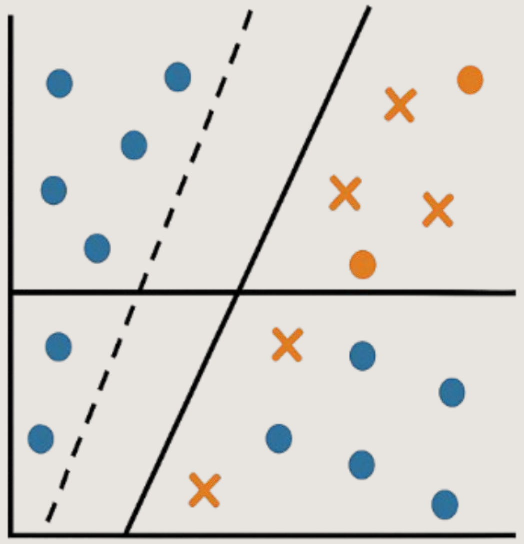
Boxplots and intensity statistics

Models Evaluated

Logistic Regression

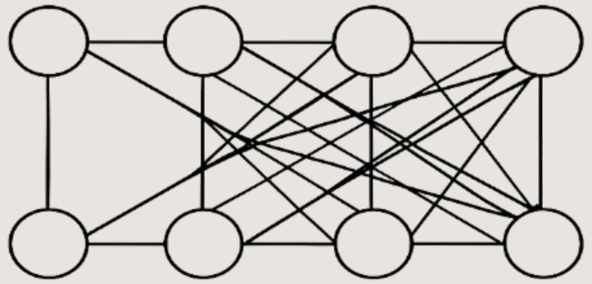


SVM with Polynomial Kernel



Neural Network

(1 hidden layer, 64 units)



Purpose

Compare linear vs. non-linear approaches.

Logistic Regression

Key Features:

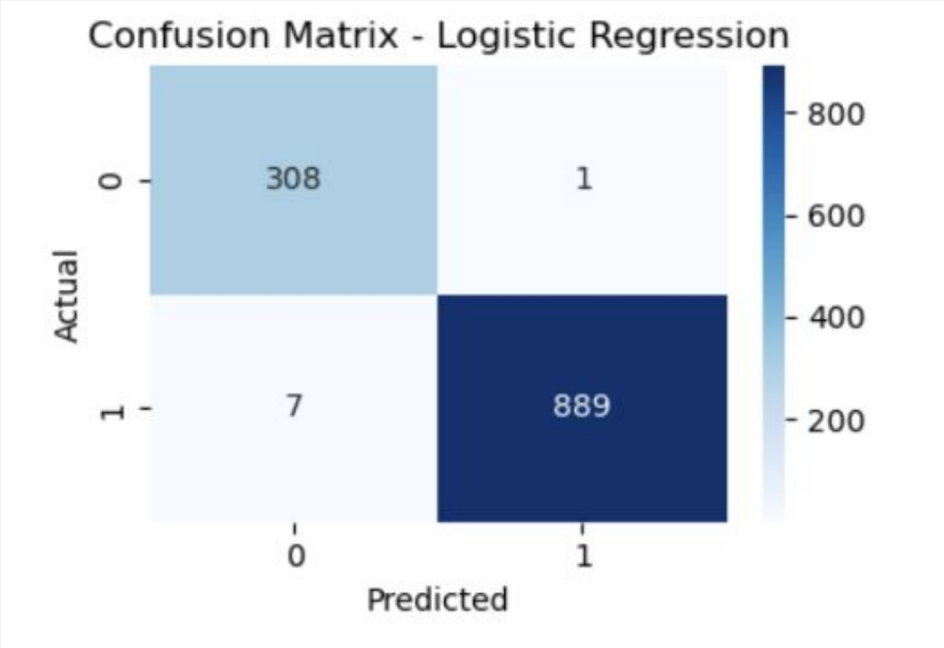
- Linear classifier with L2 regularization
- liblinear solver
- Input features:

Results:

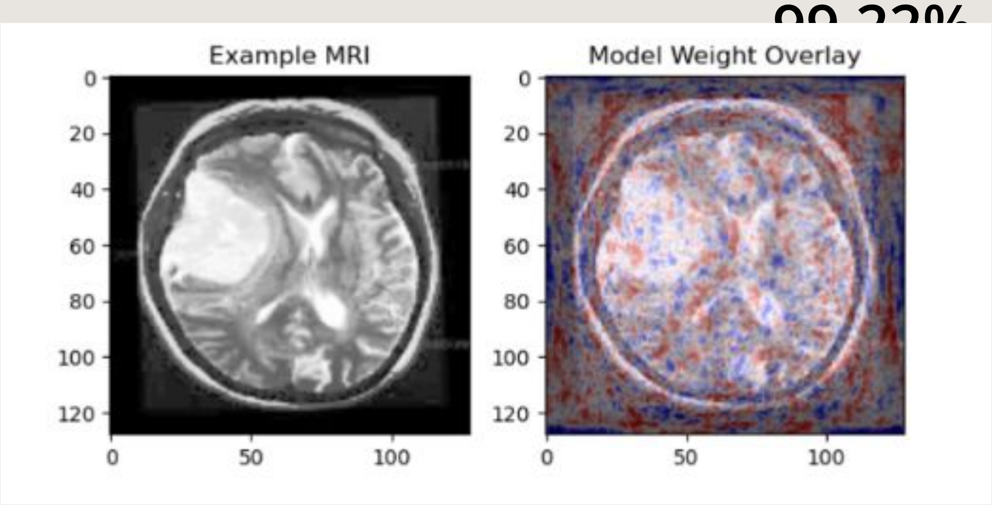
- Accuracy: 99.34%
- Precision: 99.89%
- Recall:

Include:

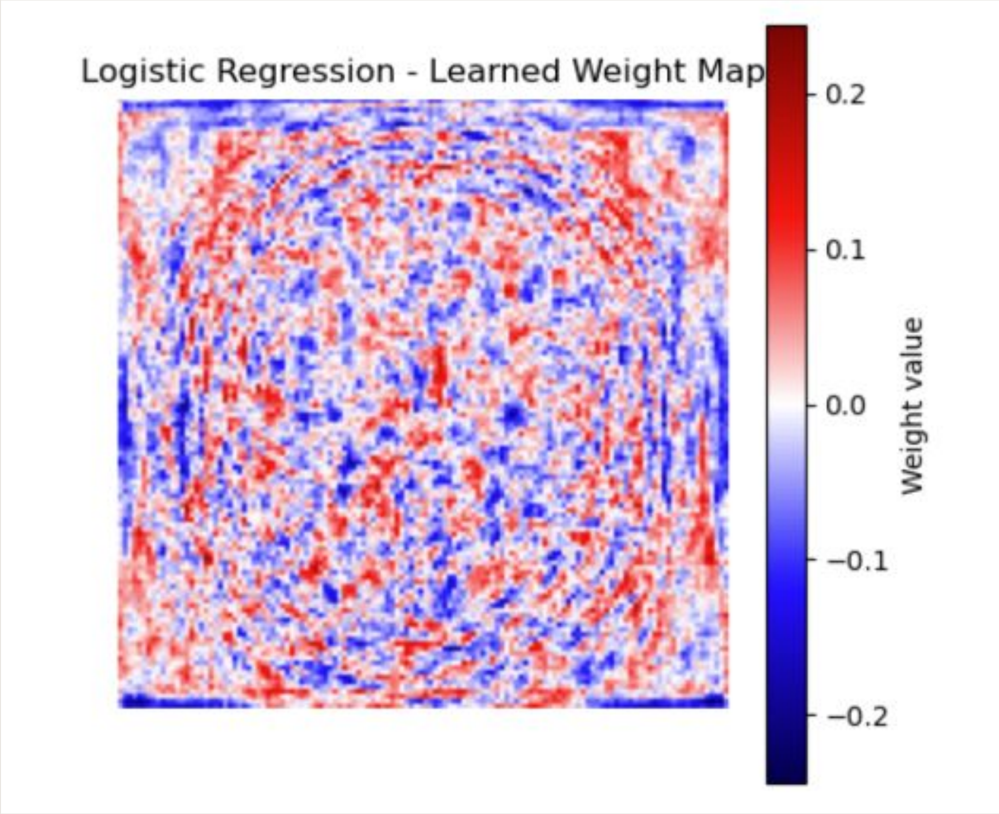
- Confusion matrix
- Weight map & overlay visualization
- Regularization impact chart



Confusion Matrix



Example MRI vs Model Weight Overlay

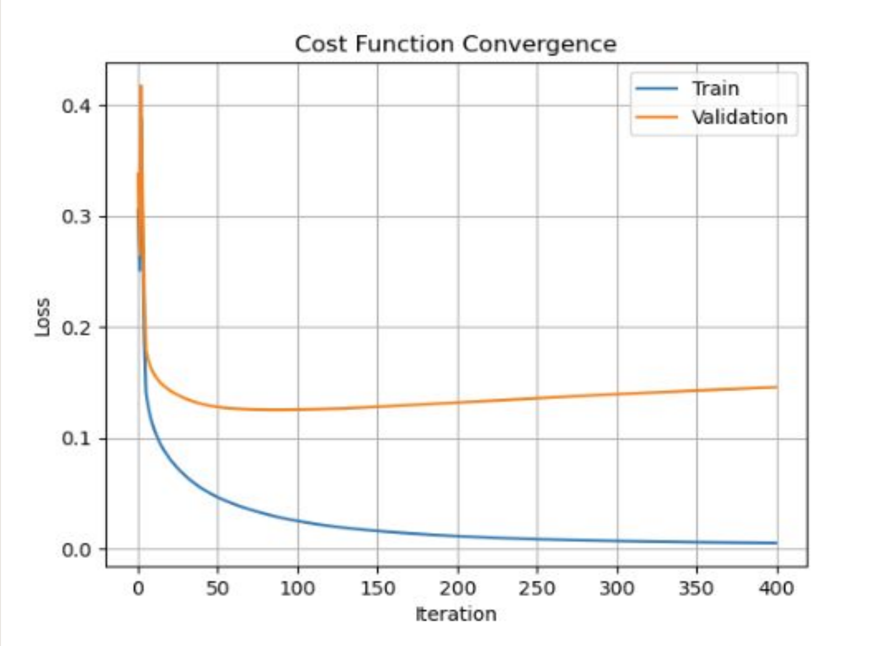


Learned Weight Map

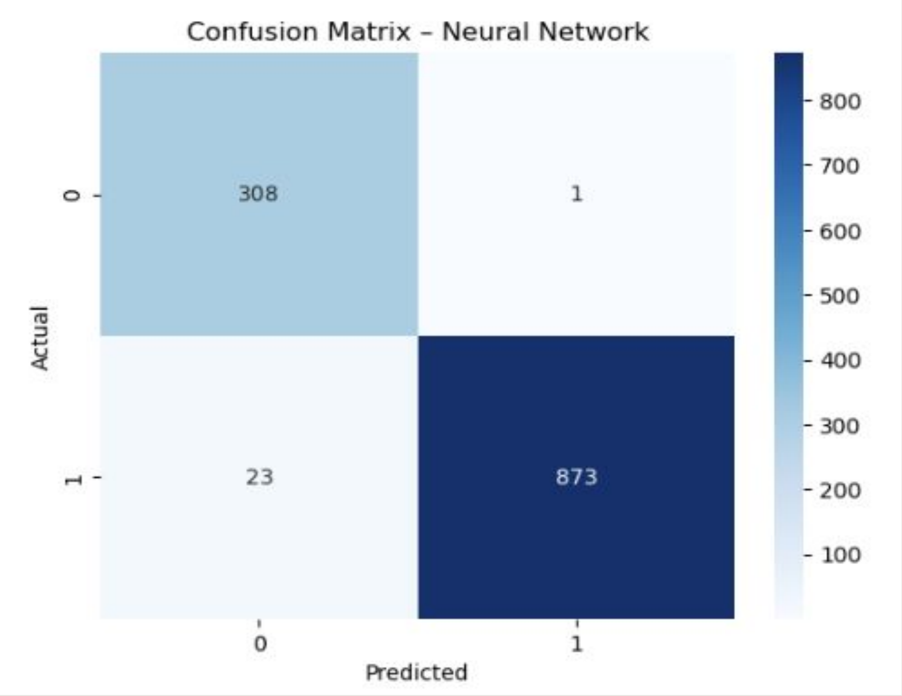
Neural Networks

- **Architecture**
  - Input layer: 16,384 features (flattened 128×128 MRI)
  - Hidden layer: 64 neurons, sigmoid activation
  - Output layer: 1 sigmoid neuron (tumor probability)
  - Implementation: fully built from scratch in NumPy
  - Optimization: batch gradient descent, L2 regularization

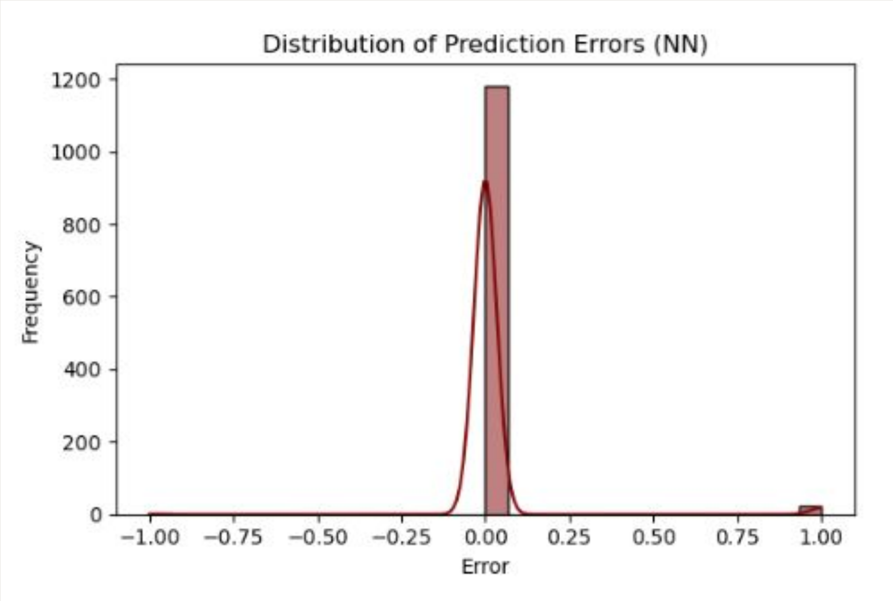
- **Hyperparameters**
  - Learning rate $\alpha = 0.5$
  - Regularization $\lambda = 0.1$
  - Iterations: 400
  - Weight initialization: Xavier



Training and validation loss



Confusion Matrix



Distribution of prediction errors

SVM

Test

Tested Linear, RBF, Polynomial kernels

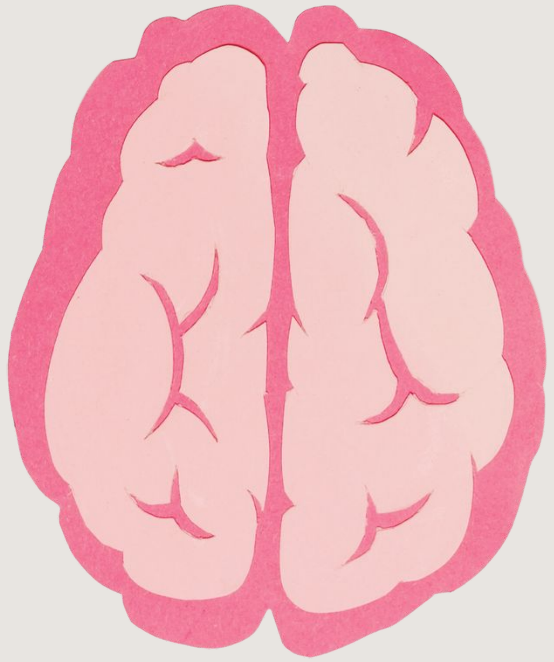
Best

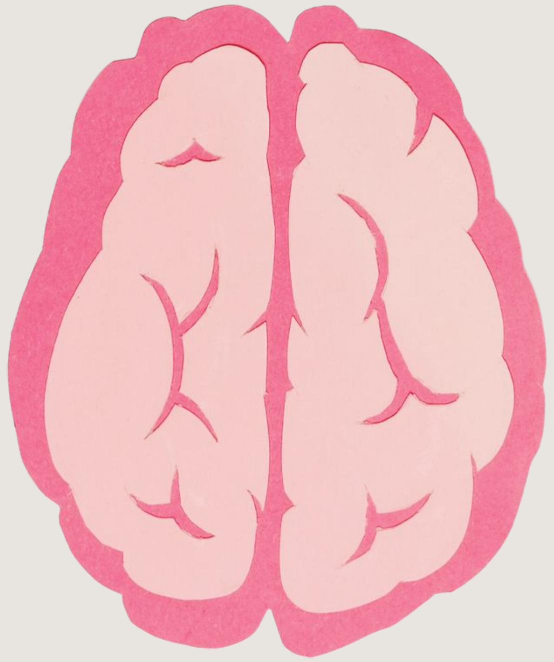
Polynomial (degree 2, C=1, γ=0.01)

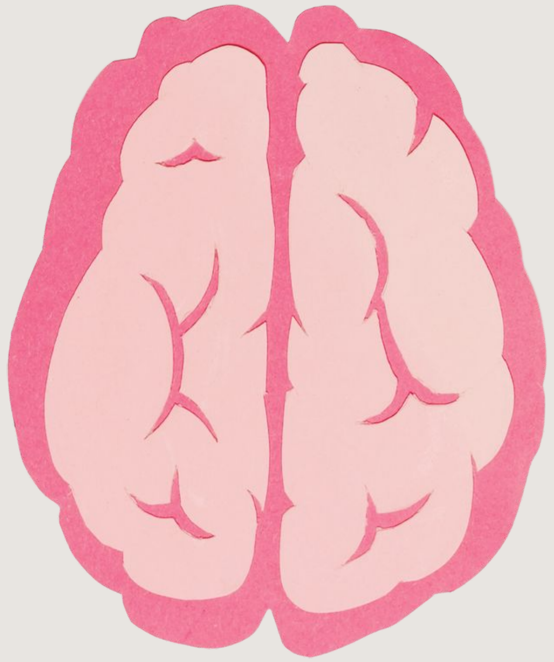
Accuracy

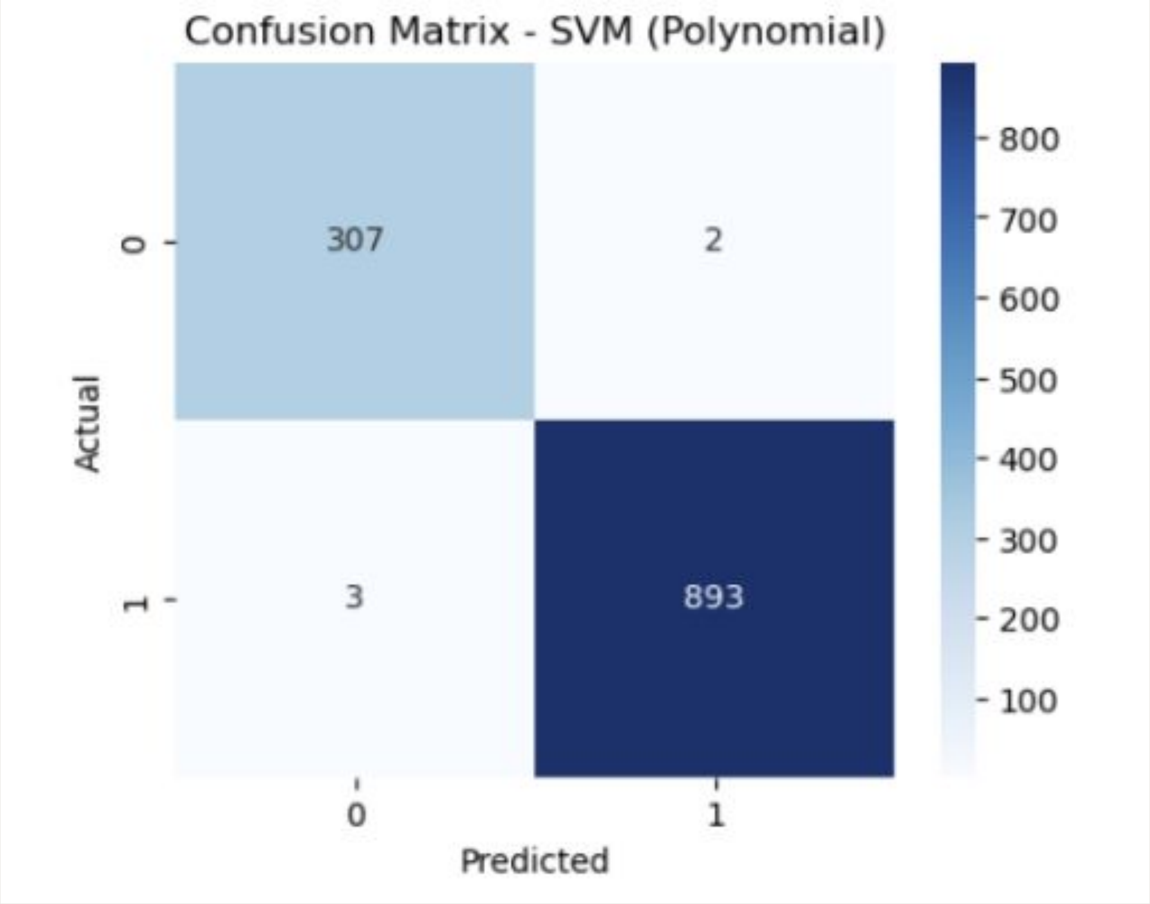
Highest: 99.59%

Strong generalization + low error rate.

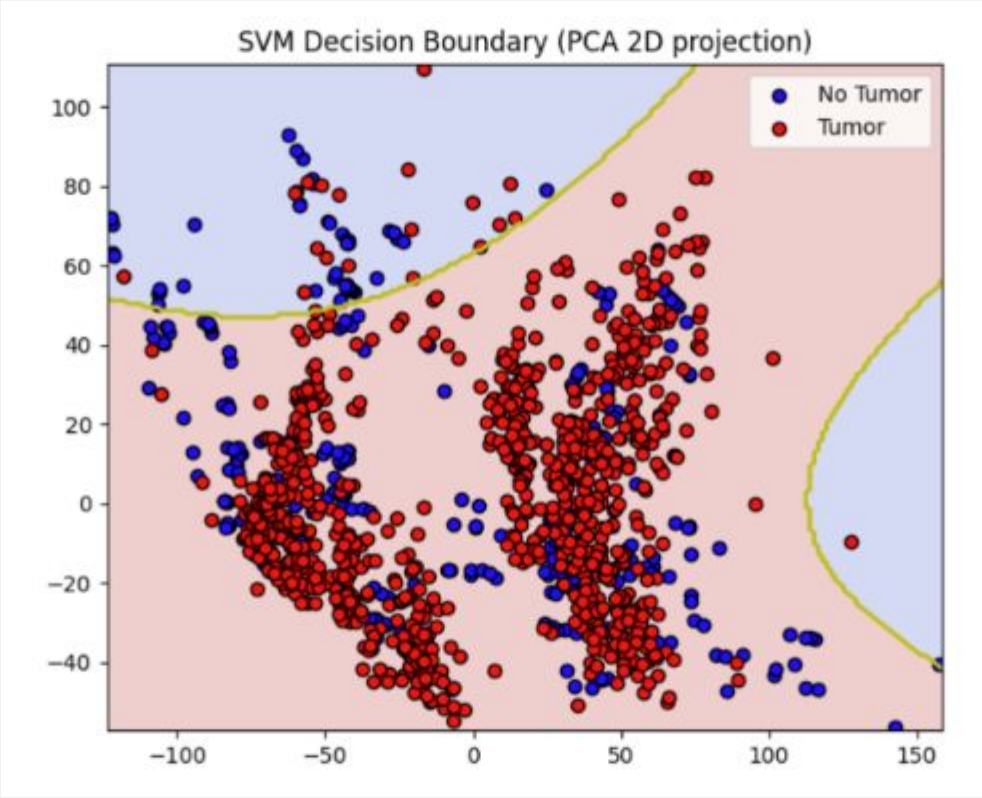




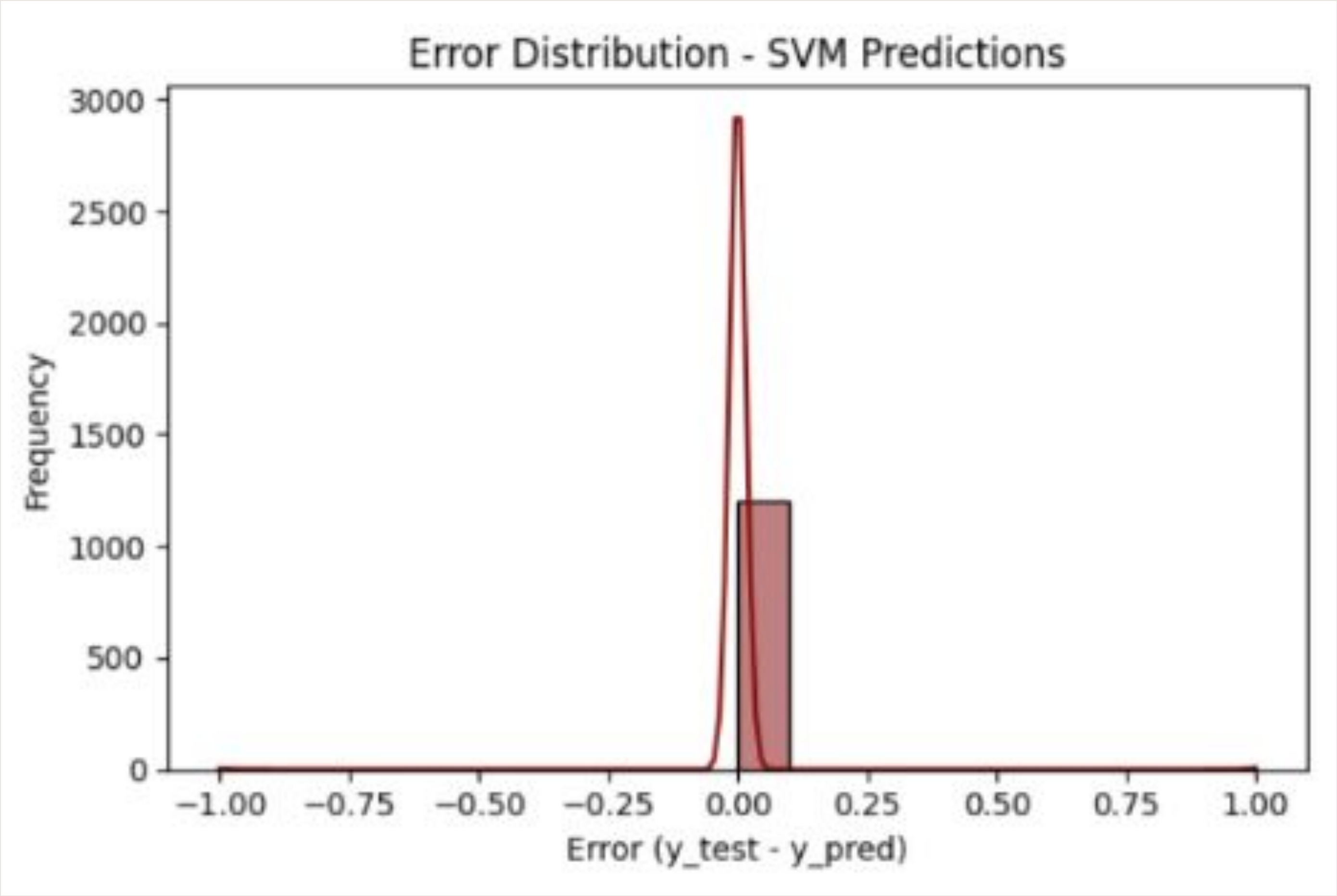




Confusion Matrix

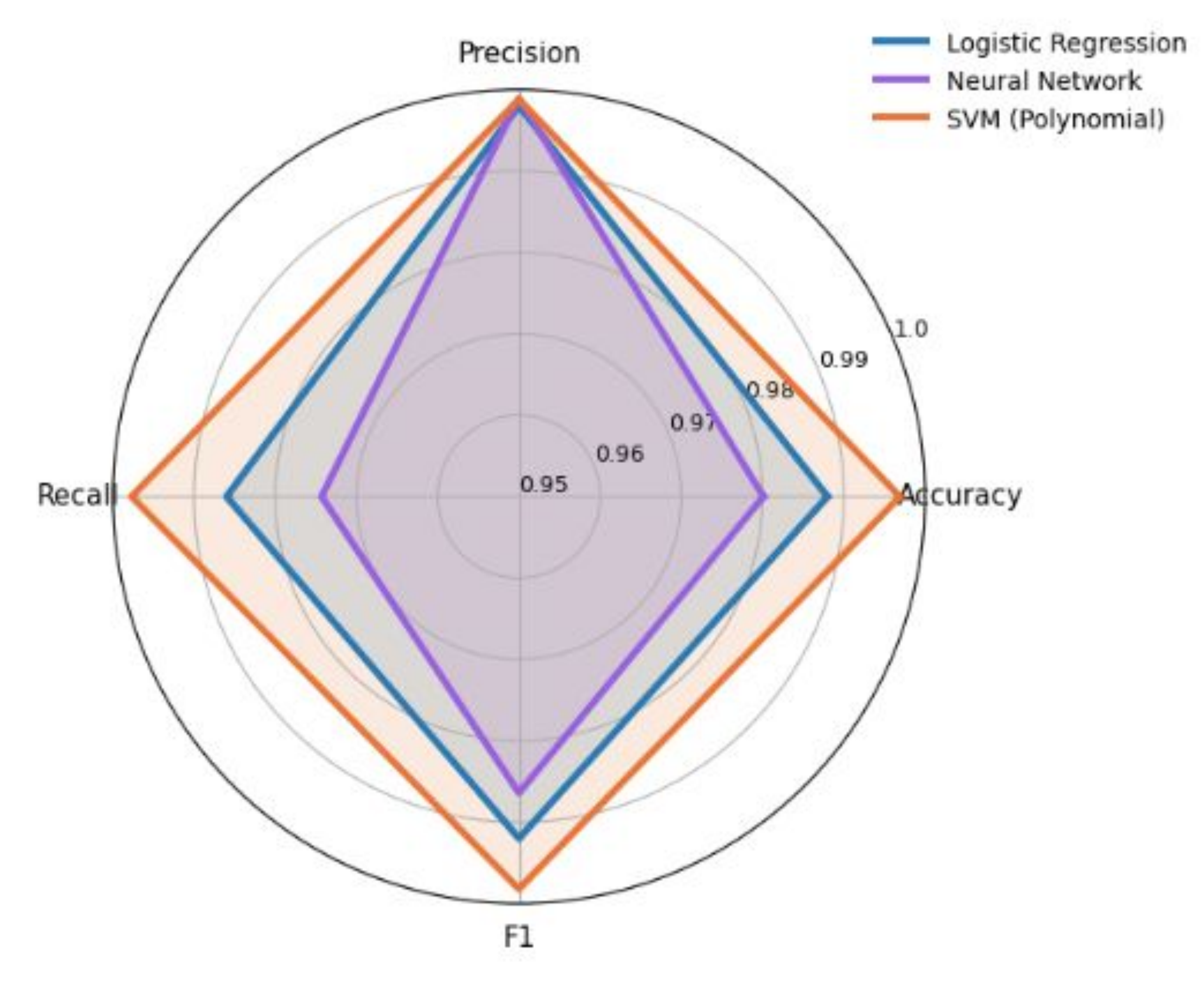


SVM Decision Boundary



Error distribution

Analysis of results



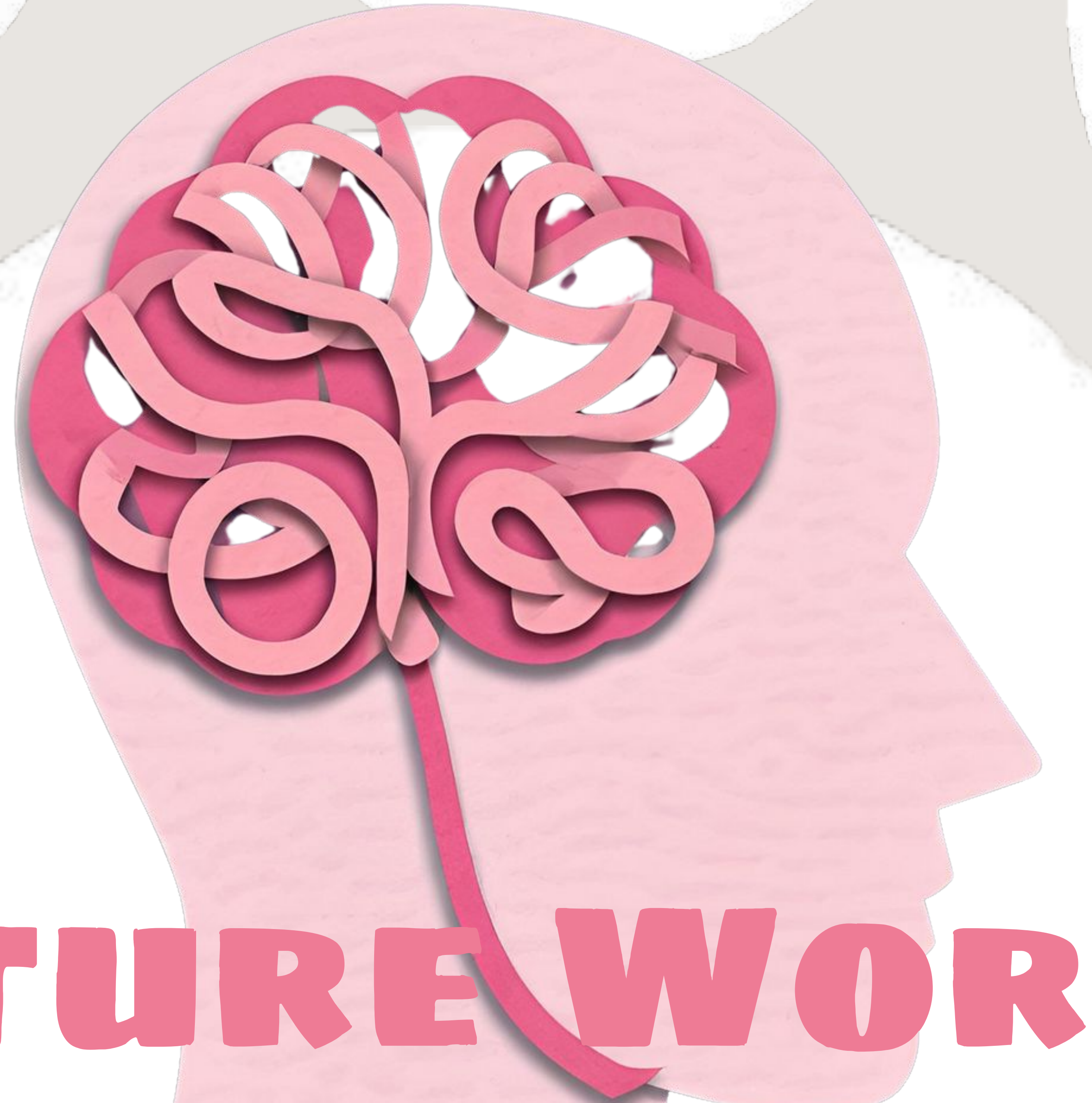
01 All models >98% accuracy

02 SVM > Logistic Regression > Neural Network

03 NN affected by overfitting (higher FN)

04 Pre-processing crucial for separability

Future Work



A

Extend to multi-class tumor identification

B

Use CNNs or transfer learning

C

Apply XAI (Grad-CAM, LIME)

D

Evaluate on multi-institution datasets

E

Improve NN regularization

WORK DIVISION

Carolina Silva	• Conclusion • Neural Network • State of the Art • Analysis of Results
Matilde Teixeira	• Introduction • Pre-processing • SVM • Logistic Regression

References

1. National Brain Tumor Society – Brain Tumor Facts & Awareness
2. IARC / WHO – CNS Cancer Fact Sheet (GLOBOCAN 2022)
3. Nickparvar, *Brain Tumor MRI Dataset*, Kaggle (2021)
4. Elgohary, *GLCM + ML for Brain Tumor Detection* (2024)
5. Basthikodi et al., *SVM for Multiclass Brain Tumor Diagnosis* (2024)
6. Khan et al., *CNN for Brain Tumor Classification*, MBE (2020)
7. Sajjad et al., *ML-based Detection of Brain Tumours*, BSPC (2022)
8. Saxena et al., *Hybrid CNN–SVM Framework*, CBM (2023)
9. Hemanth & Patil, *CNN–SVM Hybrid Models*, IEEE Access (2023)
10. Sun et al., *Data Augmentation & Transfer Learning*, Neurocomputing (2023)
11. Mittal et al., *XAI in Neuro-Oncology*, CSBJ (2025)
12. Dorfner, *Deep Learning for Brain MRI Tumor Analysis*, NPJ Precision Oncology (2025)
13. Liv Hospital – Brain Tumor MRI Dataset List (2025)
14. Goodfellow et al., *Deep Learning* (MIT Press, 2016)
15. Litjens et al., *DL in Medical Imaging Survey*, MedIA (2017)
16. Pan & Yang, *Transfer Learning Survey*, IEEE TKDE (2010)
17. Adadi & Berrada, *XAI Survey*, IEEE Access (2018)

THANKS!
QUESTIONS?