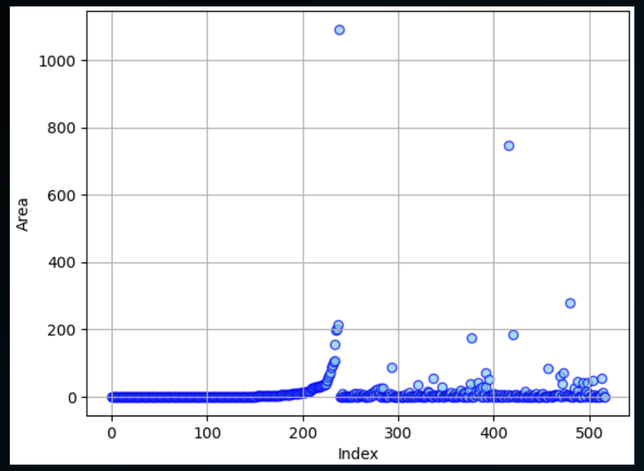
EMBERITHMS

WILDFIRE RECOGNITION AND IMPACT ASSESSMENT

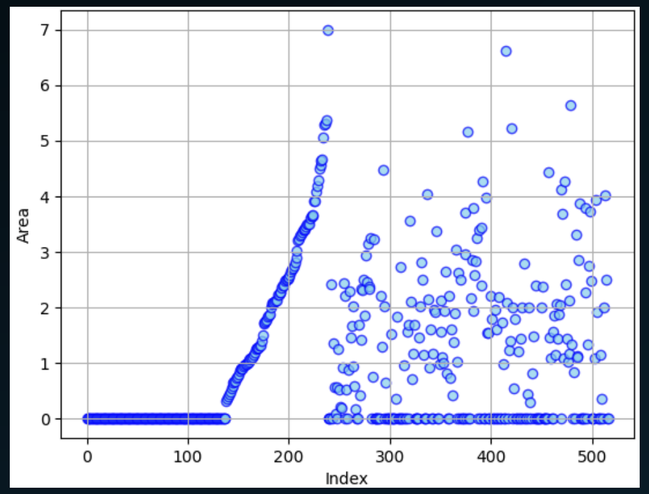
DATA MINING & MACHINE LEARNING

Adil Alizada, Aditi Singh, Johan Thomas,
Nazeem Ahmed, Rachael Dias

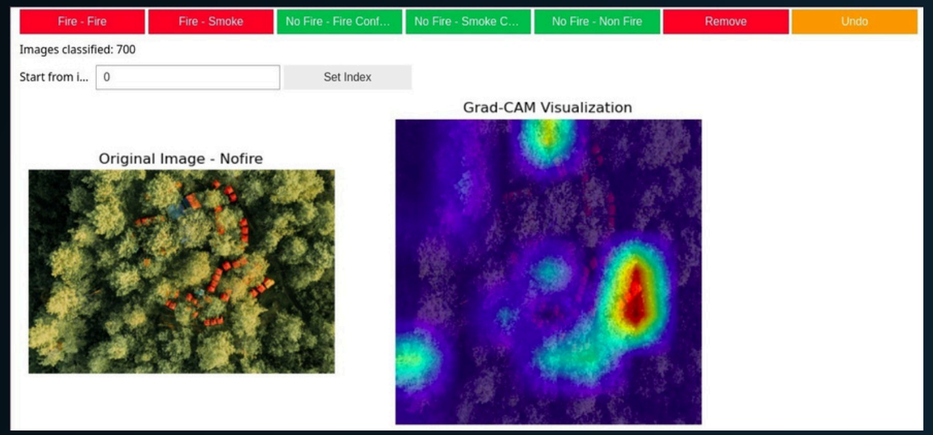
WORKING WITH OUR DATASETS



NUMERICAL DATASET BEFORE LOG TRANSFORMATION



AFTER LOG TRANSFORMATION



INTERFACE FOR IMAGE RECLASSIFICATION

NUMERICAL DATASET DECISIONS

- REMOVED DAY FEATURE (DID NOT CORRESPOND TO TARGET VARIABLE)
- LABEL ENCODING FOR MONTH FEATURE

UNDERSAMPLED TO THE MINORITY CLASS BEFORE ONLINE AUGMENTATION

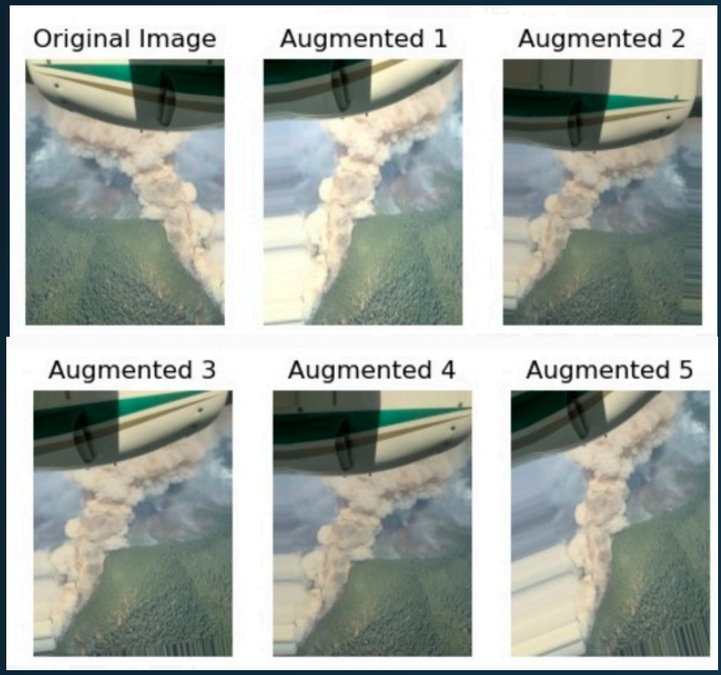
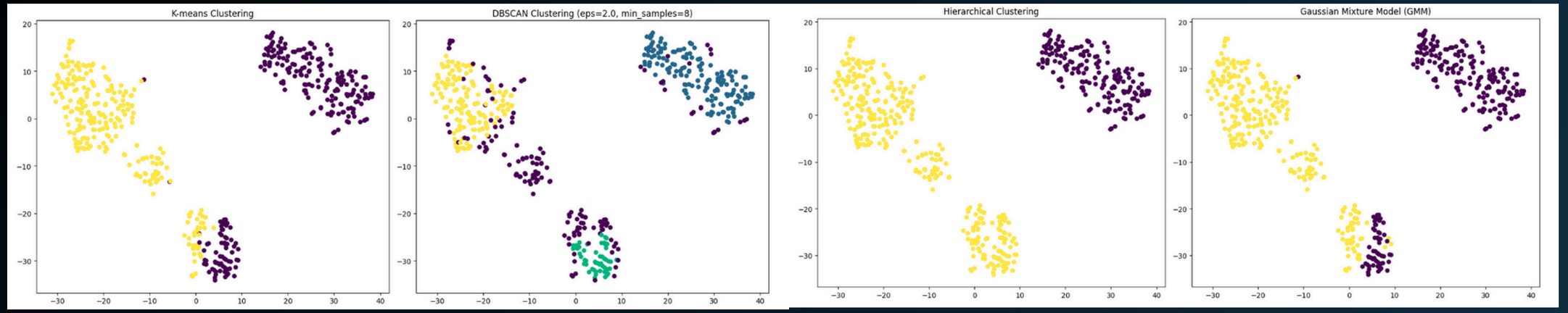
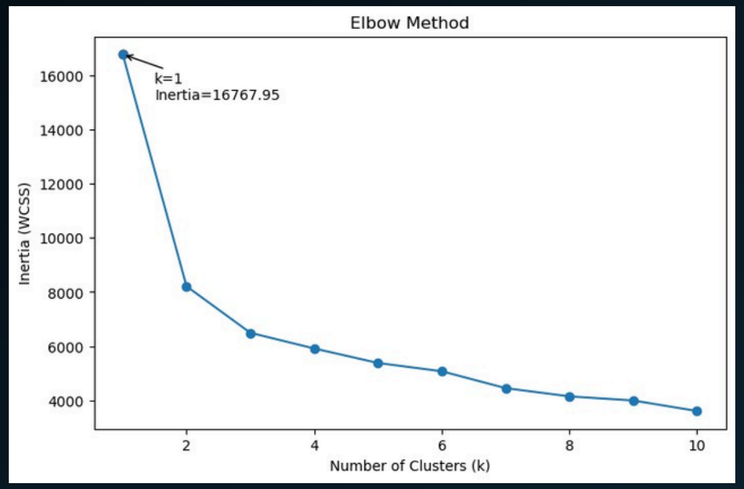


IMAGE AUGMENTATION EXAMPLES

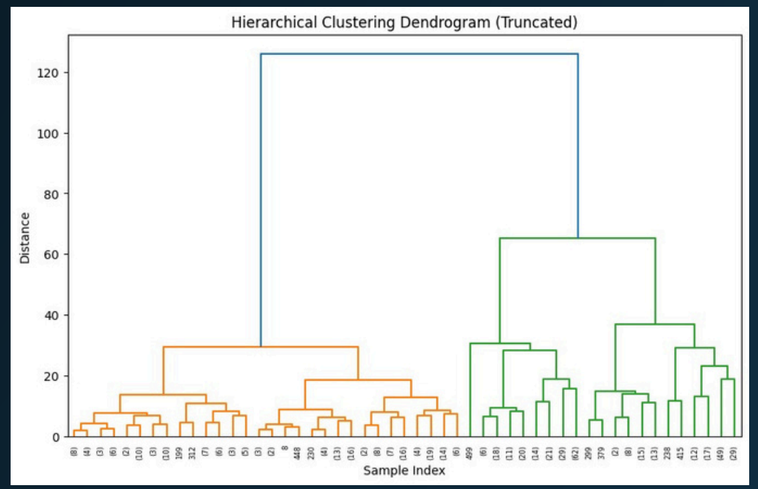
CLUSTERING



t-SNE VISUALISATION



ELBOW PLOT

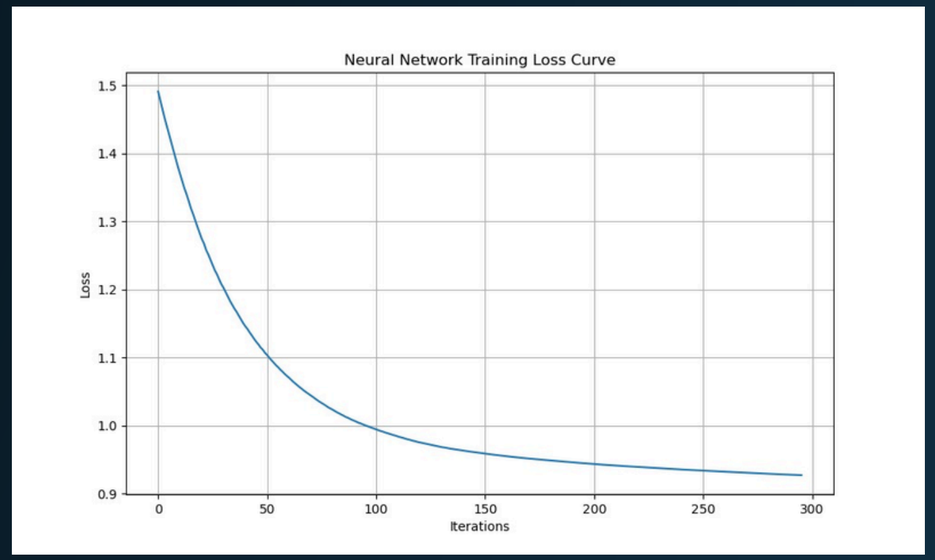


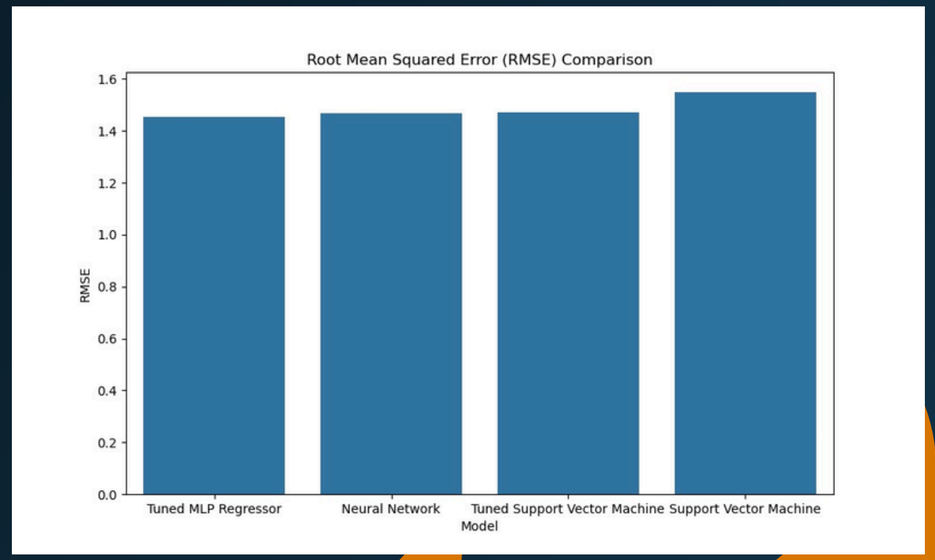
CLUSTERING DENDOGRAM

REGRESSION AND CLASSIFICATION MODELS

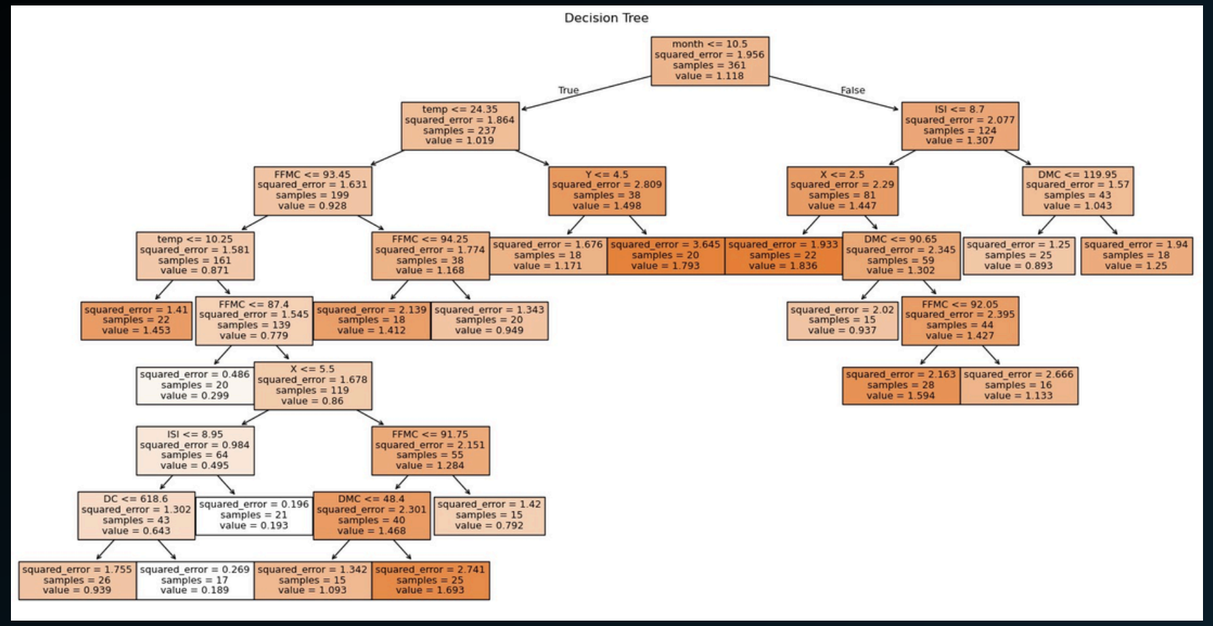
Model	MSE
Linear Regression	2.06
kNN Regression	1.914
Random Forest	1.919
MLP	2.10
SVM	2.16

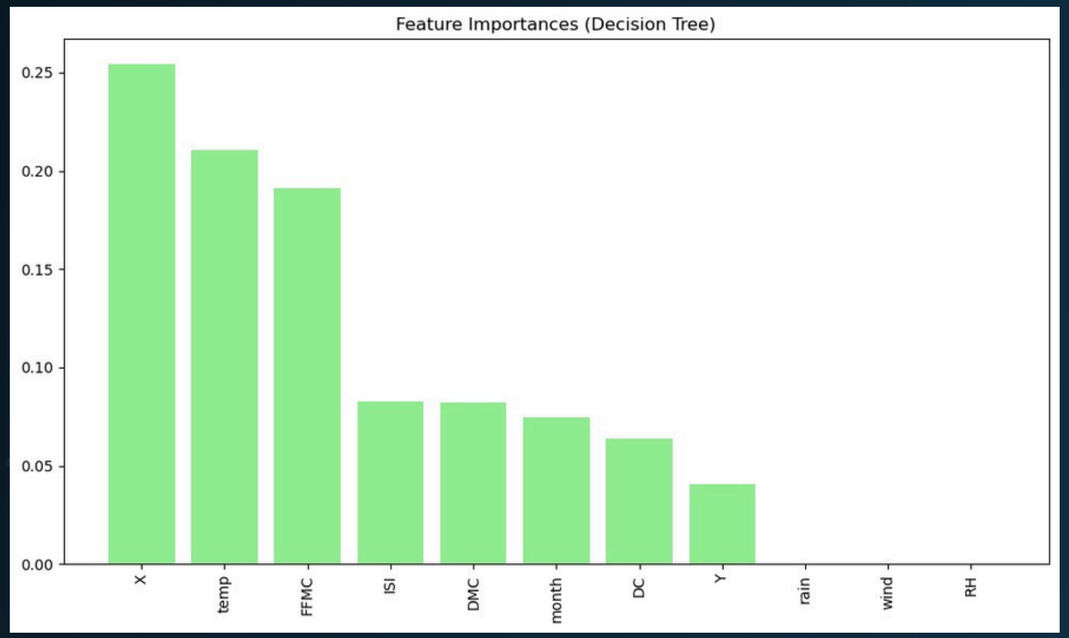
Model (Without Area class 0)	Accuracy	Precision	Recall
Multinomial NB	51.85%	26.89	51.85%
Gaussian NB	18.5%	0.03%	18.5%
Categorical NB	41.98%	40.39%	41.98%
Bernoulli NB	51.85%	39.83	51.85%





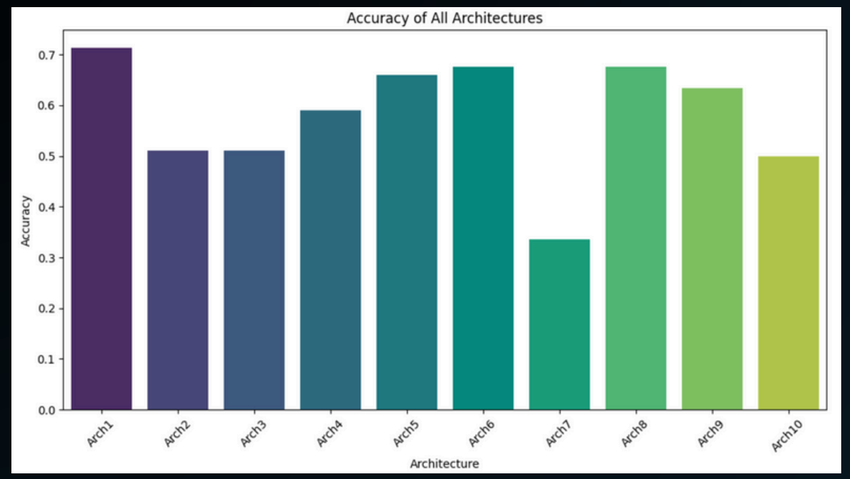
DECISION TREES



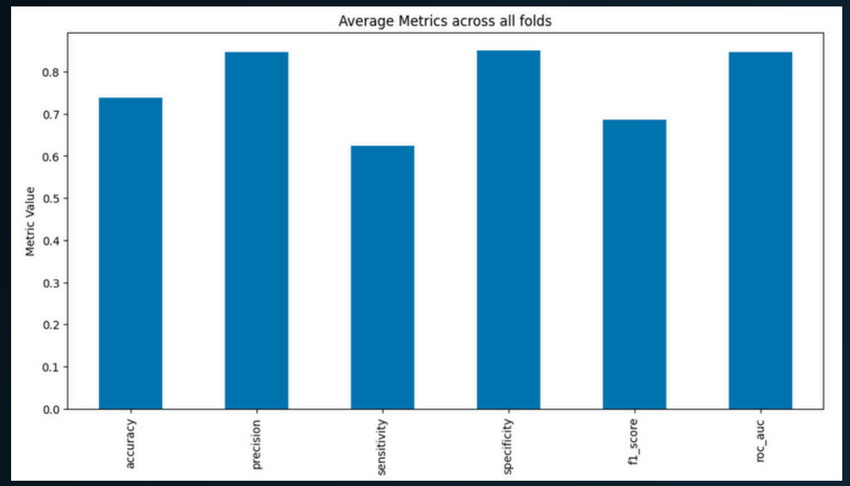


- HYPER-PARAMETER TUNING USING OPTUNA
- TRAIN-TEST SPLIT WITH DIFFERENT RATIOS

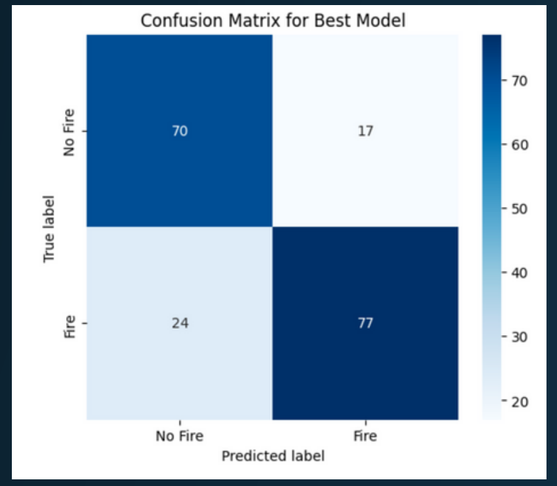
CONVOLUTIONAL NEURAL NEWORKS



ACCURACY OF 10 UNIQUE ARCHITECTURES



10 FOLD VALIDATION ON BEST MODEL



CONFUSION MATRIX FOR BEST MODEL

10 UNIQUE ARCHITECTURES WERE MADE, EACH VARYING IN:

- IN THE NUMBER OF CONVOLUTIONAL LAYERS (1-4)
- NUMBER OF FILTERS (32-256)
- KERNEL SIZE (3,5)
- FULLY CONNECTED UNITS (64-256)
- AND DROPOUT RATE (0.2-0.5)

- HYPERPARAMETER TUNING WAS CONDUCTED VARYING THE LEARNING RATE, DROPOUT RATE AND BATCH SIZES.
- INCLUDING KFOLD VALIDATION AND HYPERPARAMETER TUNING, A TOTAL OF 35 CNN MODELS WERE TRAINED.