Name: Sushruth Danivasa Sridhar

Assignment – 3

Task 1 (Regression):

1. A brief description of the dataset.

For this regression task, I have chosen dataset from the OpenML database with dataset ID 550. The regression task involves predicting a continuous target variable based on several input features. The dataset consists of numerous features which include both numerical and categorical data, capturing diverse aspects of the underlying problem domain. This dataset contains 2178 number of instances and 4 features. The data was collected from a house located in Stam Bruges, Belgium, and includes various attributes related to the house's temperature and humidity conditions in different rooms, weather observations from a nearby weather station, and random variables that are indicative of energy management practices. The goal is to predict the amount of energy used by appliances, which is a continuous variable, making it a suitable candidate for regression analysis.

2. Description of the Four Models Tried.

Model 1: Simple Model

- Consists of two dense layers with 64 and 32 neurons, respectively, and uses the ReLU activation function.

Model 2: Increased Layers

- This model has three dense layers with 128, 64, and 32 neurons, adding complexity to capture more patterns in the data. This model also uses ReLU activation function.

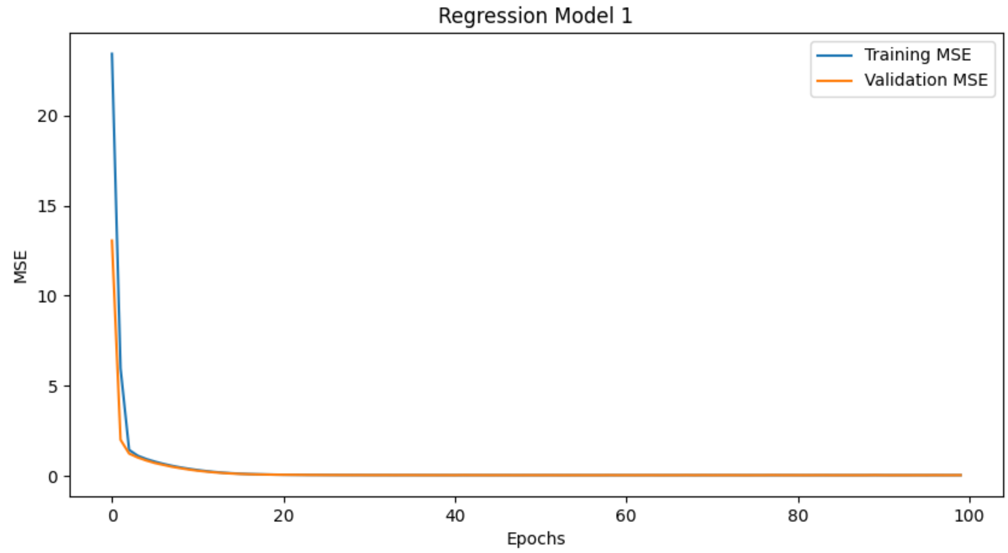
Model 3: Increased Nodes, Different activation function and Dropout

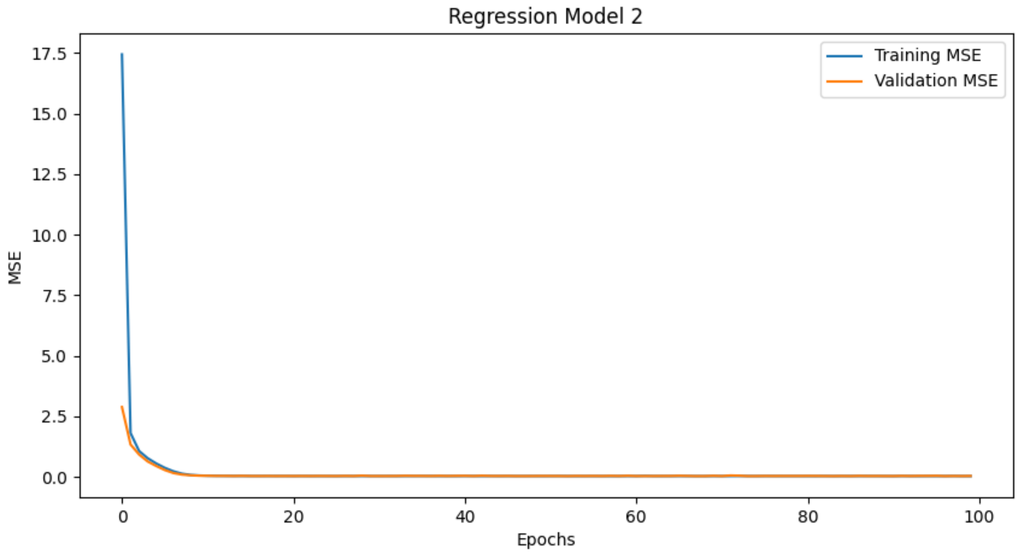
- This model has one ReLU and sigmoid activation function. Includes more neurons 256 and 128 and dropout layers to prevent overfitting, aiming for better generalization on unseen data.

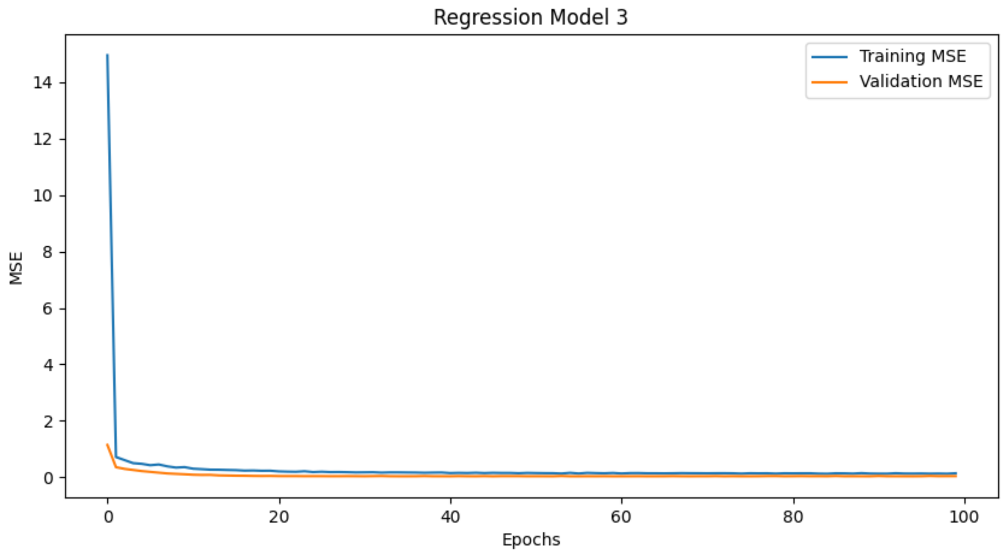
Model 4: Same Activation Functions

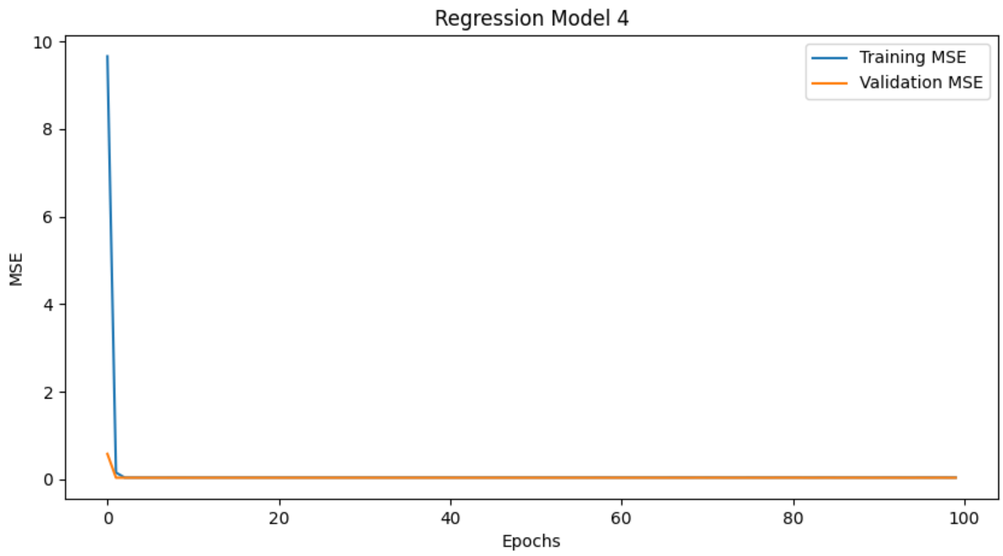
- Uses sigmoid activation functions and consists of two dense layers with 128 and 64 neurons exploring the impact of same activations on regression performance.

3. Four graphs, one for each model.


Regression Model 1
Training MSE
Validation MSE
MSE
Epochs


Regression Model 2
Training MSE
Validation MSE
MSE
Epochs


Regression Model 3
Training MSE
Validation MSE
MSE
Epochs



4. A table of minimum validation errors.

Model	Minimum Validation MSE
1	0.036457
2	0.037709
3	0.033988
4	0.034035

5. Discussion of the results.

The regression analysis across four different models demonstrated a quick convergence and closely matched training and validation MSE, indicating effective generalization without overfitting. Model 1, the simplest configuration with two dense layers, managed to deliver a satisfactory performance, showing that a basic structure has the potential to capture the essential patterns in the data. Model 2 introduced more layers, yet the increase in complexity did not translate to a significant improvement in the validation MSE, which remained in close proximity to that of Model 1. Model 3 yielded the lowest validation MSE, suggesting a slight edge in capturing the dataset's patterns, potentially due to its architecture. Model 4 also shows promise, despite its simpler structure compared to Model 3, pointing out that sometimes less complex architectures can be nearly as effective. Therefore, Model 3 is marginally superior.

Task 2 (Classification):

1. A brief description of the dataset.

For this classification task, I have chosen dataset from the OpenML database with dataset ID 1462. The banknote-authentication dataset is utilized for a binary classification problem, which entails predicting one of two possible discrete labels based on the input features. This dataset comprises statistical attributes derived from image data, such as variance, skewness, kurtosis, and entropy of wavelet transformed images. The dataset includes 1372 instances, each described by 4 features. The features do not represent raw image or textual data, complying with restrictions against using such data directly. Instead, they provide a transformed representation of the images, suitable for machine learning tasks that require pattern recognition without processing pixel data. Variance (V1) measure of the spread of the pixel intensity values. Skewness (V2) measure of the asymmetry of the pixel intensity distribution. Curtosis (V3) measure of the 'tailedness' of the pixel intensity distribution, indicating the presence of outliers. Entropy (V4) measure of the randomness or texture of the image, indicating the complexity of the image content. The target for the classification task is a binary variable indicating one of two possible outcomes, 1 for genuine and 2 for forged.

2. Description of the Four Models Tried.

Model 1: Simple Model

- It consists of two dense layers with 64 and 32 neurons, one using the sigmoid and another one using the ReLU activation function respectively. The final output layer uses a sigmoid activation to produce a probability indicating the likelihood of the instance belonging to the positive class.

Model 2: Same Activation Functions

- This model uses sigmoid activations throughout, aiming to explore the impact of consistent sigmoid activations in hidden layers. It consists of a single dense layer with 64 neurons followed by the output layer with a sigmoid activation, focusing on simplicity and continuity in activation functions.

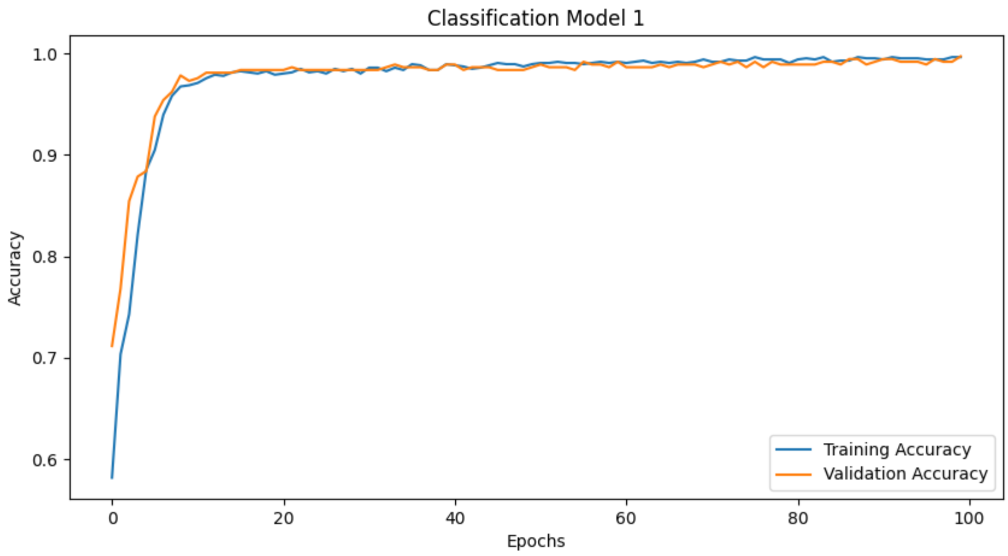
Model 3: Same Activations, more dense layers and Dropout

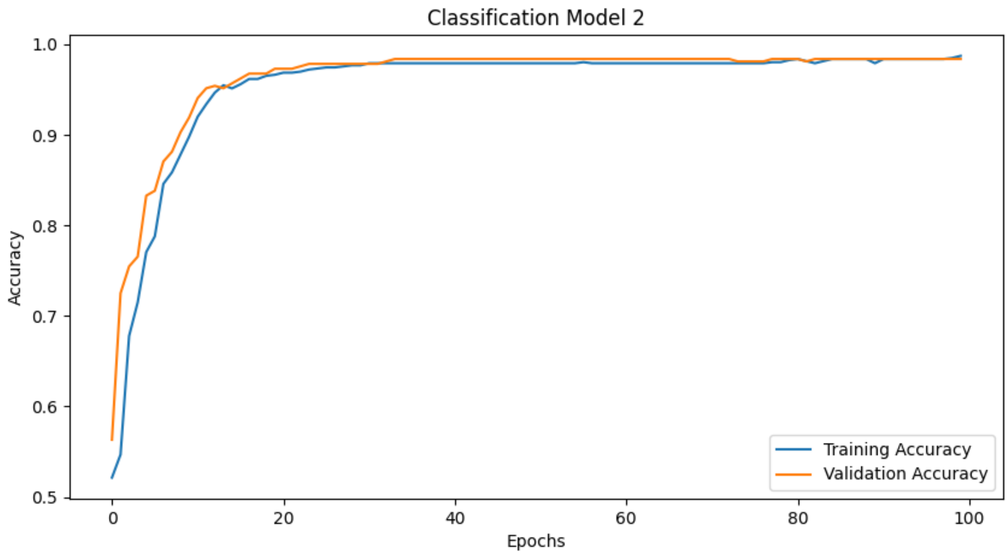
- Same activation functions and an increased number of neurons. It starts with a dense layer of 512 neurons using the sigmoid activation, followed by layers with 256 and 128 neurons using sigmoid activations, respectively. Dropout layers are used to prevent overfitting. The final layer uses a sigmoid activation to predict the class.

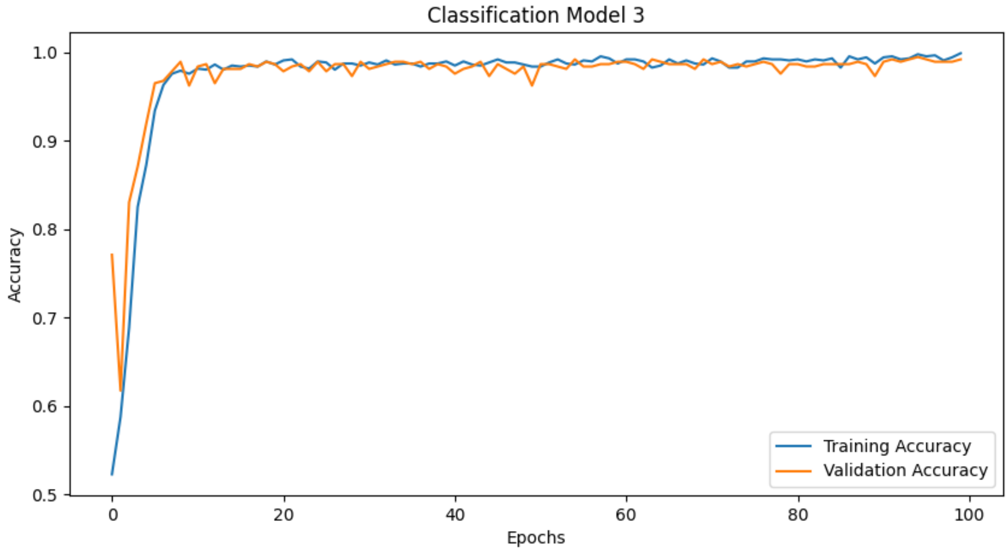
Model 4: Mixed Activation Functions

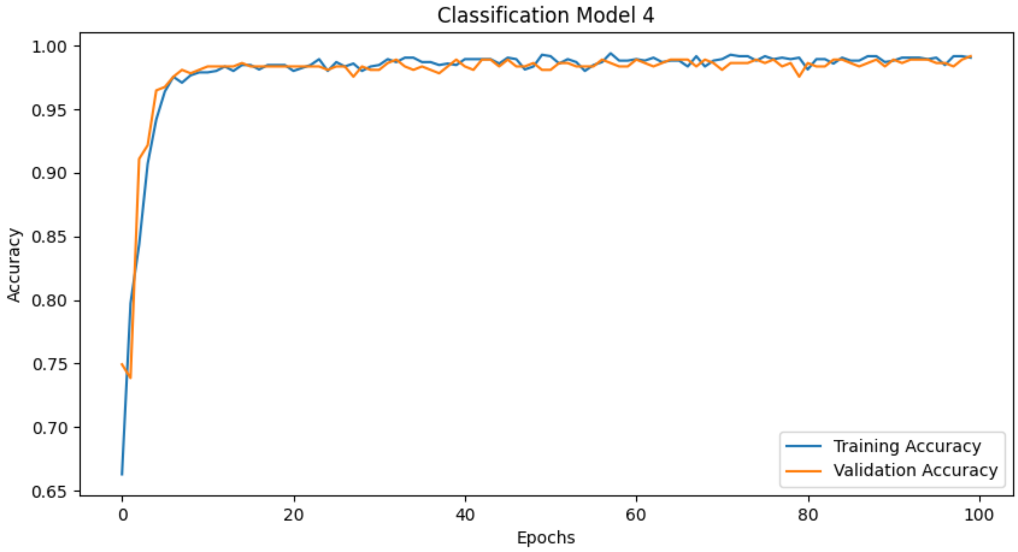
- It starts with a dense layer of 256 neurons using the sigmoid activation, followed by a layer of 128 neurons with ReLU activation, and concludes with a sigmoid-activated output layer.

3. Four graphs, one for each model.


Classification Model 1
Accuracy
Epochs
Training Accuracy
Validation Accuracy


Classification Model 2
Accuracy
Epochs
Training Accuracy
Validation Accuracy


Classification Model 3
Accuracy
Epochs
Training Accuracy
Validation Accuracy



4. A table of maximum validation accuracies.

Model	Maximum Validation Accuracy
1	0.997305
2	0.983827
3	0.994609
4	0.991914

5. Discussion of the results.

For the classification task, the models achieved high validation accuracy, all above 98%, which is indicative of their strong predictive performance. Model 1 reached near-perfect accuracy, showing that even a simple architecture could effectively capture the patterns in the dataset. However, the slight differences in maximum validation accuracy suggest that increased complexity in models, such as additional layers or neurons, does not necessarily translate to significantly improved performance on this dataset. The results from Models 2 to 4, with various configurations, converged to similar accuracy levels, suggesting that the dataset characteristics may be such that simpler models are sufficient to achieve high accuracy, and additional complexity does not yield a proportional gain in predictive power.