Zillow Prize: Zillow's Home Value Prediction (Zestimate)

The goal for the notebook is predicting whether the house price is above or below median value using ann.

Description of notebook: The first part of the notebook is preparing the data for the model. The second part is regarding setting up the neural network with set hyperparameters, training the model and visualizing the results. And lastly, the third part is hypertuning the model to see if the model will perform better with different hyperparameters.

1. Processing the data

1.1. Importing needed imports

In [1]: import pandas as pd

1.2. Retrieving the data

Data explanation can be found below for each column:

- Lot Area (in sq ft)
- Overall Quality (scale from 1 to 10)
- Overall Condition (scale from 1 to 10)
- Total Basement Area (in sq ft)
- Number of Full Bathrooms
- Number of Half Bathrooms
- Number of Bedrooms above ground
- Total Number of Rooms above ground
- Number of Fireplaces
- Garage Area (in sq ft)

The last column of the dataset is what we would like to predict: AboveMedianPrice. For example: Is the house price above the median or not? (1 for yes and 2 for no)

```
df = pd.read_csv('housepricedata.csv')
In [3]:
         df.head()
            LotArea OverallQual OverallCond TotalBsmtSF FullBath HalfBath BedroomAbvGr TotRmsAbvGrd Fireplaces GarageArea AboveMedianPrice
Out[3]:
                                          5
                                                                                        3
                                                                                                       8
         0
              8450
                             7
                                                     856
                                                               2
                                                                                                                 0
                                                                                                                           548
                                                                                                                                               1
                                          8
                                                               2
                                                                         0
         1
              9600
                                                   1262
                                                                                        3
                                                                                                                 1
                                                                                                                           460
                                                                                                                                               1
         2
              11250
                             7
                                          5
                                                     920
                                                               2
                                                                                        3
                                                                                                       6
                                                                                                                 1
                                                                                                                           608
                                                                                                                                               1
         3
              9550
                             7
                                                    756
                                                                         0
                                                                                                       7
                                                                                                                 1
                                                                                                                           642
                                                                                                                                               0
         4
                             8
                                                   1145
                                                               2
                                                                                        4
                                                                                                       9
                                                                                                                 1
                                                                                                                           836
              14260
                                                                         1
                                                                                                                                               1
```

1.3. Converting the data for our machine to process

we will be converting the dataframe to an array for the machine to be able to process the data.

```
In [5]:
        dataset = df.values
        dataset
Out[5]: array([[ 8450,
                                    5, ...,
                                                     548,
                                                               1],
                                   8, ...,
                [ 9600,
                                                     460,
                                                               1],
                [11250,
                                                     608,
                                                               1],
                [ 9042,
                                   9, ...,
                                                     252,
                                                               1],
                            7,
                [ 9717,
                                   6, ...,
                                                     240,
                                                               0],
                                                               0]], dtype=int64)
                [ 9937,
                                   6, ...,
                                                     276,
```

1.4. Splitting the features and prediction data

In this section we will be splitting the data into the input and output data (X & Y).

```
In [6]: X = dataset[:, 0:10]
In [7]: Y = dataset[:, -1]
```

1.5. Scaling the feautres

In this section we will be scaling the input data to have similiar input value. For example not having one input being over 1000 and another input is in a range of 5. Due to the amount being so big, will make it difficault for the initialization of the neural network.

1.5.1. Importing needed imports

```
In [8]: from sklearn import preprocessing
```

1.5.2. Scaling the data

We will be using the min-max scaler which scales the input data so that the input features are between the values 0 and 1. Scaling down to values between 0 & 1 will help aid the training of the neural network.

```
In [9]: min_max_scaler = preprocessing.MinMaxScaler()
        X scale = min max scaler.fit transform(X)
In [10]: X scale
Out[10]: array([[0.0334198 , 0.66666667, 0.5 , ..., 0.5 , 0.
                0.3864598 ],
                                              , ..., 0.33333333, 0.33333333,
               [0.03879502, 0.55555556, 0.875
               0.32440056],
               [0.04650728, 0.66666667, 0.5 , ..., 0.33333333, 0.33333333,
               0.42877292],
               [0.03618687, 0.66666667, 1. , ..., 0.58333333, 0.66666667,
               0.17771509],
                                             , ..., 0.25 , 0. ,
               [0.03934189, 0.44444444, 0.625
               0.16925247],
                                             , ..., 0.33333333, 0. ,
               [0.04037019, 0.44444444, 0.625
                0.19464034]])
```

1.6. Splitting the data into a train, test & validation sets

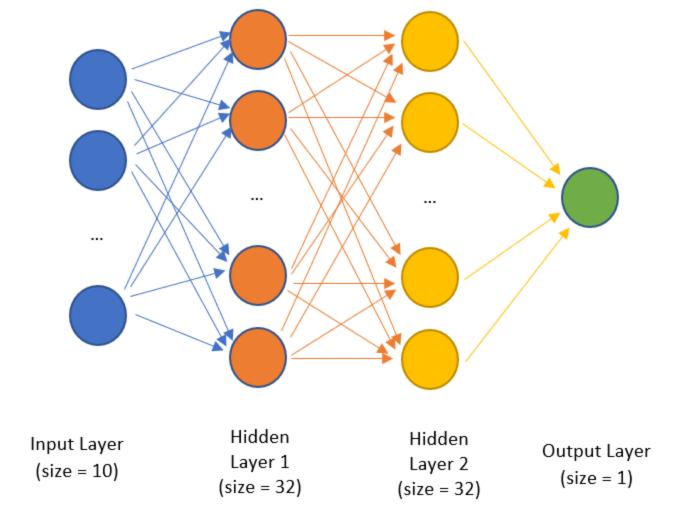
1.6.1. Importing needed imports

```
In [11]: from sklearn.model_selection import train_test_split
```

1.6.2. Splitting the data

2. Building & Training the Neural Network

2.1. Setting up the architecture



We will be setting the layers as followed:

- Hidden layer 1: 32 neurons, ReLU activation
- Hidden layer 2: 32 neurons, ReLU activation
- Output Layer: 1 neuron, Sigmoid activation

Different activation function that can be applied:

- Sigmoid Function (A function which 'squeezes' all the initial output to be between 0 and 1)
- tanh Function (A function which 'squeezes' all the initial output to be between -1 and 1)
- ReLU Function (If the initial output is negative, then output 0. If not, do nothing to the initial output)

```
from keras.layers import Dense
```

We will be setting the model layers

2.2. Configuring & Training the Model

The architecture for the model has been set but we still need to configure the model. These configurations include:

- what optimizer will be used such as SGD & Adam.
- what lost function will be used? (first checking if it is for Probabilistic or regression losses) such as BinaryCrossentropy & CategoricalCressentropy.
- what other metrics will like to be tracker such as Accuray & MeanSquaredError.

In this section we will be conducting the followed steps:

- 1. Specify some hyper-parameters (the template)
- 2. Train on the training dataset (filling in the parameters)
- 3. Record the validation loss
- 4. Applying confusion matrix on the output
- 5. conclusion

Time to train the model

```
Epoch 1/100
Epoch 2/100
Epoch 3/100
Epoch 4/100
Epoch 5/100
Epoch 6/100
Epoch 7/100
Epoch 8/100
Epoch 9/100
Epoch 10/100
Epoch 11/100
Epoch 12/100
Epoch 13/100
Epoch 14/100
Epoch 15/100
Epoch 16/100
Epoch 17/100
Epoch 18/100
Epoch 19/100
Epoch 20/100
Epoch 21/100
Epoch 22/100
Epoch 23/100
Epoch 24/100
32/32 [=======================] - 0s 4ms/step - loss: 0.5306 - accuracy: 0.8611 - val loss: 0.5336 - val accuracy: 0.8539
```

```
Epoch 25/100
Epoch 26/100
Epoch 27/100
Epoch 28/100
Epoch 29/100
Epoch 30/100
Epoch 31/100
Epoch 32/100
Epoch 33/100
Epoch 34/100
Epoch 35/100
Epoch 36/100
Epoch 37/100
Epoch 38/100
Epoch 39/100
Epoch 40/100
Epoch 41/100
Epoch 42/100
Epoch 43/100
Epoch 44/100
Epoch 45/100
Epoch 46/100
Epoch 47/100
Epoch 48/100
32/32 [=======================] - 0s 4ms/step - loss: 0.3520 - accuracy: 0.8708 - val loss: 0.3705 - val accuracy: 0.8539
```

```
Epoch 49/100
Epoch 50/100
Epoch 51/100
Epoch 52/100
Epoch 53/100
Epoch 54/100
Epoch 55/100
Epoch 56/100
Epoch 57/100
Epoch 58/100
Epoch 59/100
Epoch 60/100
Epoch 61/100
Epoch 62/100
Epoch 63/100
Epoch 64/100
Epoch 65/100
Epoch 66/100
Epoch 67/100
Epoch 68/100
Epoch 69/100
Epoch 70/100
Epoch 71/100
Epoch 72/100
32/32 [=======================] - 0s 5ms/step - loss: 0.2966 - accuracy: 0.8836 - val loss: 0.3231 - val accuracy: 0.8630
```

```
Epoch 73/100
Epoch 74/100
Epoch 75/100
Epoch 76/100
Epoch 77/100
Epoch 78/100
Epoch 79/100
Epoch 80/100
Epoch 81/100
Epoch 82/100
Epoch 83/100
Epoch 84/100
Epoch 85/100
Epoch 86/100
Epoch 87/100
Epoch 88/100
Epoch 89/100
Epoch 90/100
Epoch 91/100
Epoch 92/100
Epoch 93/100
Epoch 94/100
Epoch 95/100
Epoch 96/100
32/32 [=======================] - 0s 4ms/step - loss: 0.2751 - accuracy: 0.8894 - val loss: 0.3013 - val accuracy: 0.8721
```

Evaluating the model

Summary of the model

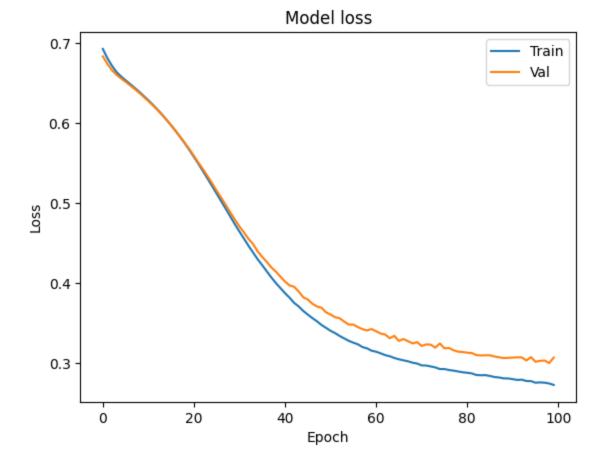
- setting up the architecture with keras. This included the input layer, 2 hidden layers and 1 output layer.
- compiling the model by seeting the settings of the model (optimizer, loss function & metrics)
- training the model and finding the best fit parameters with using the validation data.
- evaluating the nodel on the test dataset.

2.3. Visualizing Loss and Accuracy

This chapter will be focussed on visualizing the training loss and validation loss to see how our model trained. We will be looking how well our model performs and if we have overfitted. Overfitting occurs when the model has fitted so well to the training dataset that it has failed to generalize to unseen examples. This is characterized by a high dev (validation) loss and a low train loss, and can be addressed with regularization techniques. These techniques can be see as adding regularization, early stopping or dropout.

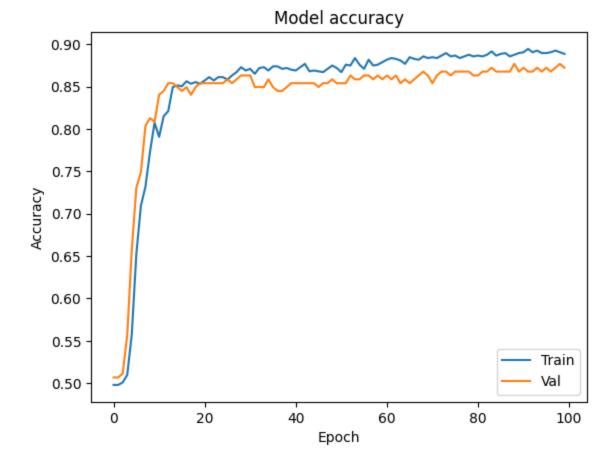
```
In [20]: # Needed imports
import matplotlib.pyplot as plt

In [21]: plt.plot(hist.history['loss'])
   plt.plot(hist.history['val_loss'])
   plt.title('Model loss')
   plt.ylabel('Loss')
   plt.xlabel('Epoch')
   plt.legend(['Train', 'Val'], loc='upper right')
   plt.show()
```



Next we will be plotting the training accuracy and validation accuracy

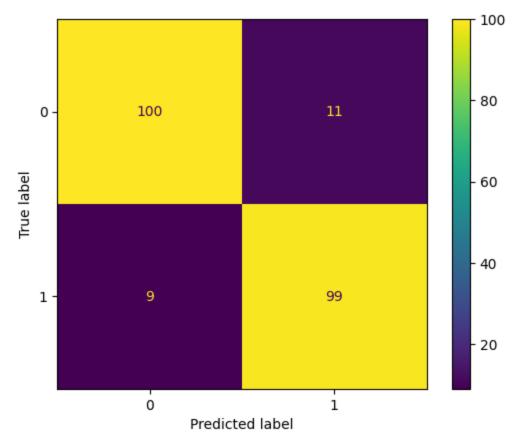
```
In [22]: plt.plot(hist.history['accuracy'])
    plt.plot(hist.history['val_accuracy'])
    plt.title('Model accuracy')
    plt.ylabel('Accuracy')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='lower right')
    plt.show()
```



the plot above shows the model's accuracy for the training and validation set

2.4 Confusion Matrix

```
In [26]: cm = confusion_matrix(Y_test, y_pred)
    disp = ConfusionMatrixDisplay(confusion_matrix=cm)
    disp.plot()
    plt.show()
```



0 1	0.92 0.90	0.90 0.92	0.91 0.91	111 108
accuracy			0.91	219
macro avg	0.91	0.91	0.91	219
weighted avg	0.91	0.91	0.91	219

3. Hypertuning the model

In this section we will be hypertuning the optimizer and loss function for the model to see if it will perform better or worse and why.

3.1. Hypertuning the model optimizers

In this chapter we will use the same steps for setting up the neural network but play around with the hyper paramters to see if we can get a better model by evaluating the loss and accuracy.

In this section we will be training and hypertuning the model to find the best results. The steps will go as followed:

- 1. Specify some hyper-parameters (the template)
- 2. Train on the training dataset (filling in the parameters)
- 3. Applying confusion matrix on the output
- 4. Record the validation loss
- 5. Repeat Steps 1 to 3 with a different set of hyper-parameters (many times)
- 6. Conclusion

3.1.1 Using Adam as optimizer

Adam optimization is a stochastic gradient descent method that is based on adaptive estimation of first-order and second-order moments.

Neural Network Architecture

Training the Model

```
Epoch 1/100
Epoch 2/100
Epoch 3/100
Epoch 4/100
Epoch 5/100
Epoch 6/100
Epoch 7/100
Epoch 8/100
Epoch 9/100
Epoch 10/100
Epoch 11/100
Epoch 12/100
Epoch 13/100
Epoch 14/100
Epoch 15/100
Epoch 16/100
Epoch 17/100
Epoch 18/100
Epoch 19/100
Epoch 20/100
Epoch 21/100
Epoch 22/100
Epoch 23/100
Epoch 24/100
32/32 [=======================] - 0s 4ms/step - loss: 0.2451 - accuracy: 0.9041 - val loss: 0.2799 - val accuracy: 0.8767
```

```
Epoch 25/100
Epoch 26/100
Epoch 27/100
Epoch 28/100
Epoch 29/100
Epoch 30/100
Epoch 31/100
Epoch 32/100
Epoch 33/100
Epoch 34/100
Epoch 35/100
Epoch 36/100
Epoch 37/100
Epoch 38/100
Epoch 39/100
Epoch 40/100
Epoch 41/100
Epoch 42/100
Epoch 43/100
Epoch 44/100
Epoch 45/100
Epoch 46/100
Epoch 47/100
Epoch 48/100
32/32 [=======================] - 0s 4ms/step - loss: 0.2235 - accuracy: 0.9061 - val loss: 0.2500 - val accuracy: 0.8813
```

```
Epoch 49/100
Epoch 50/100
Epoch 51/100
Epoch 52/100
Epoch 53/100
Epoch 54/100
Epoch 55/100
Epoch 56/100
Epoch 57/100
Epoch 58/100
Epoch 59/100
Epoch 60/100
Epoch 61/100
Epoch 62/100
Epoch 63/100
Epoch 64/100
Epoch 65/100
Epoch 66/100
Epoch 67/100
Epoch 68/100
Epoch 69/100
Epoch 70/100
Epoch 71/100
Epoch 72/100
32/32 [=======================] - 0s 5ms/step - loss: 0.2065 - accuracy: 0.9100 - val_loss: 0.2488 - val accuracy: 0.8858
```

```
Epoch 73/100
Epoch 74/100
Epoch 75/100
Epoch 76/100
Epoch 77/100
Epoch 78/100
Epoch 79/100
Epoch 80/100
Epoch 81/100
Epoch 82/100
Epoch 83/100
Epoch 84/100
Epoch 85/100
32/32 [===============] - 0s 4ms/step - loss: 0.2060 - accuracy: 0.9110 - val loss: 0.2645 - val accuracy: 0.8721
Epoch 86/100
Epoch 87/100
Epoch 88/100
Epoch 89/100
Epoch 90/100
Epoch 91/100
Epoch 92/100
Epoch 93/100
Epoch 94/100
Epoch 95/100
Epoch 96/100
```

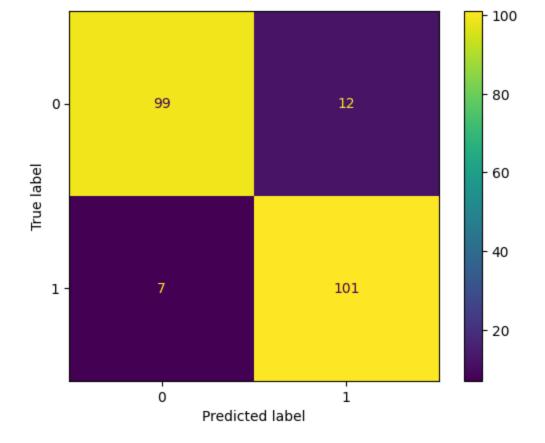
```
Epoch 97/100
Epoch 98/100
Epoch 99/100
32/32 [============== - 0s 5ms/step - loss: 0.2126 - accuracy: 0.9119 - val loss: 0.2594 - val accuracy: 0.8721
Epoch 100/100
```

Evaluating the Model

```
In [31]: model_2.evaluate(X_test, Y_test)
    Out[31]: [0.23562783002853394, 0.913241982460022]
```

Confusion Matrix

```
In [32]: Y pred = model 2.predict(X test).round()
        y pred = np.round(Y pred, 0).tolist()
         confusion matrix(Y test, y pred)
        7/7 [======== ] - 0s 2ms/step
Out[32]: array([[ 99, 12],
               [ 7, 101]], dtype=int64)
In [33]: cm = confusion_matrix(Y_test, y_pred)
        disp = ConfusionMatrixDisplay(confusion_matrix=cm)
         disp.plot()
        plt.show()
```

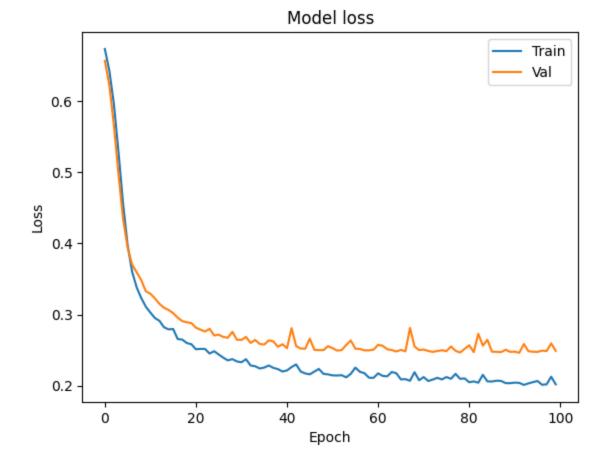


```
In [34]: print(classification_report(Y_test, y_pred))
```

	precision	recall	f1-score	support
0	0.93	0.89	0.91	111
1	0.89	0.94	0.91	108
accuracy			0.91	219
macro avg	0.91	0.91	0.91	219
weighted avg	0.91	0.91	0.91	219

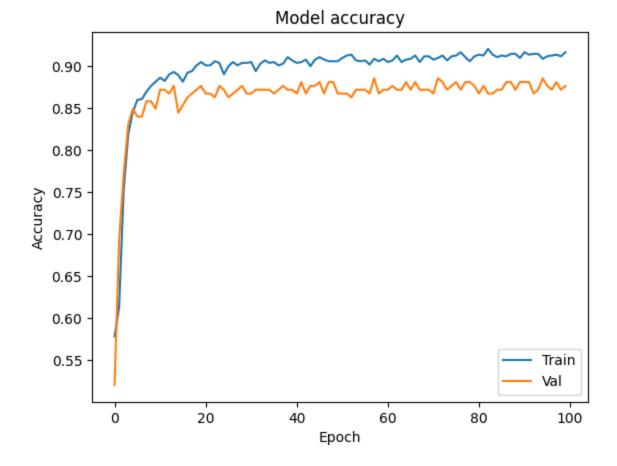
Visualizing Loss

```
In [35]: plt.plot(hist_test_2.history['loss'])
    plt.plot(hist_test_2.history['val_loss'])
    plt.title('Model loss')
    plt.ylabel('Loss')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='upper right')
    plt.show()
```



Visualizing Accuracy

```
In [36]: plt.plot(hist_test_2.history['accuracy'])
    plt.plot(hist_test_2.history['val_accuracy'])
    plt.title('Model accuracy')
    plt.ylabel('Accuracy')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='lower right')
    plt.show()
```



3.1.2 Using RMSprop as optimizer

The gist of RMSprop is to:

• Maintain a moving (discounted) average of the square of gradients

loss='binary_crossentropy', #chosen for our binary output

• Divide the gradient by the root of this average

Neural Network Architecture

```
metrics=['accuracy'])
```

Training the Model

```
Epoch 1/100
Epoch 2/100
Epoch 3/100
Epoch 4/100
Epoch 5/100
Epoch 6/100
Epoch 7/100
Epoch 8/100
Epoch 9/100
Epoch 10/100
Epoch 11/100
Epoch 12/100
Epoch 13/100
Epoch 14/100
Epoch 15/100
Epoch 16/100
Epoch 17/100
Epoch 18/100
Epoch 19/100
Epoch 20/100
Epoch 21/100
Epoch 22/100
Epoch 23/100
Epoch 24/100
32/32 [=======================] - 0s 5ms/step - loss: 0.2696 - accuracy: 0.8904 - val loss: 0.3022 - val accuracy: 0.8676
```

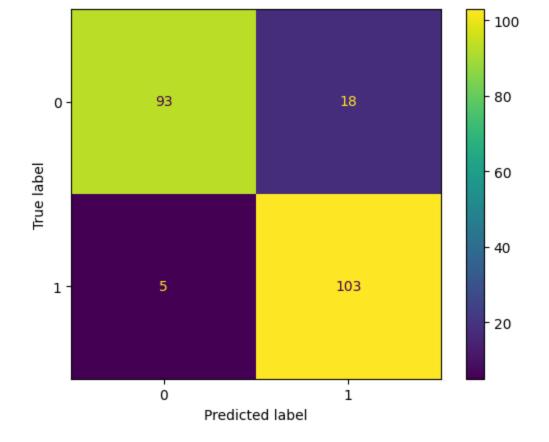
```
Epoch 25/100
Epoch 26/100
Epoch 27/100
Epoch 28/100
Epoch 29/100
Epoch 30/100
Epoch 31/100
Epoch 32/100
Epoch 33/100
Epoch 34/100
Epoch 35/100
Epoch 36/100
Epoch 37/100
Epoch 38/100
Epoch 39/100
Epoch 40/100
Epoch 41/100
Epoch 42/100
Epoch 43/100
Epoch 44/100
Epoch 45/100
Epoch 46/100
Epoch 47/100
32/32 [===============] - 0s 4ms/step - loss: 0.2319 - accuracy: 0.9022 - val loss: 0.2791 - val accuracy: 0.8767
Epoch 48/100
32/32 [=======================] - 0s 4ms/step - loss: 0.2322 - accuracy: 0.9022 - val loss: 0.2542 - val accuracy: 0.8721
```

```
Epoch 49/100
Epoch 50/100
Epoch 51/100
Epoch 52/100
Epoch 53/100
Epoch 54/100
Epoch 55/100
Epoch 56/100
Epoch 57/100
Epoch 58/100
Epoch 59/100
Epoch 60/100
Epoch 61/100
Epoch 62/100
Epoch 63/100
Epoch 64/100
Epoch 65/100
Epoch 66/100
Epoch 67/100
Epoch 68/100
Epoch 69/100
Epoch 70/100
Epoch 71/100
Epoch 72/100
32/32 [=======================] - 0s 4ms/step - loss: 0.2209 - accuracy: 0.9051 - val_loss: 0.2457 - val accuracy: 0.8721
```

```
Epoch 73/100
Epoch 74/100
Epoch 75/100
Epoch 76/100
Epoch 77/100
Epoch 78/100
Epoch 79/100
Epoch 80/100
Epoch 81/100
Epoch 82/100
Epoch 83/100
Epoch 84/100
Epoch 85/100
Epoch 86/100
Epoch 87/100
Epoch 88/100
Epoch 89/100
Epoch 90/100
Epoch 91/100
Epoch 92/100
Epoch 93/100
Epoch 94/100
Epoch 95/100
Epoch 96/100
```

Evaluating the Model

Confusion Matrix



```
In [43]: print(classification_report(Y_test, y_pred))

precision recall f1-score support

0 0.95 0.84 0.89 111
1 0.85 0.95 0.90 108
```

219

219

219

0.89

0.89

0.89

Visualizing Loss

weighted avg

accuracy macro avg

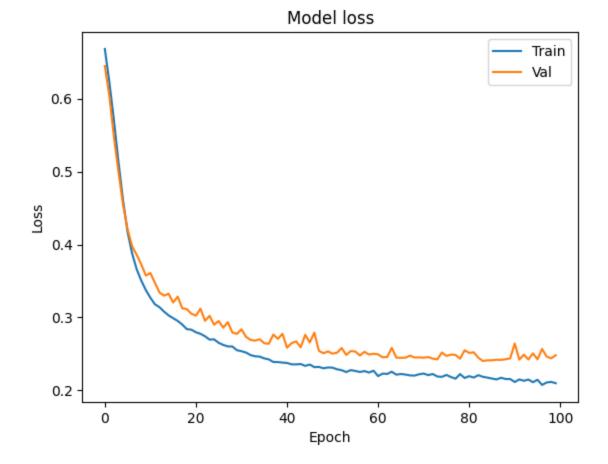
0.90

0.90

0.90

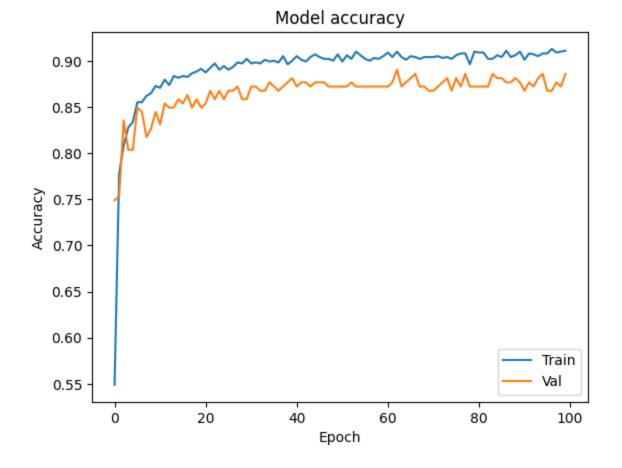
0.89

```
In [44]: plt.plot(hist_test_3.history['loss'])
    plt.plot(hist_test_3.history['val_loss'])
    plt.title('Model loss')
    plt.ylabel('Loss')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='upper right')
    plt.show()
```



Visualizing Accuracy

```
In [45]: plt.plot(hist_test_3.history['accuracy'])
    plt.plot(hist_test_3.history['val_accuracy'])
    plt.title('Model accuracy')
    plt.ylabel('Accuracy')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='lower right')
    plt.show()
```



3.1.3. Conclusion

We have tried out 2 different optimizers (Adam & RMSprop) to compare to the SGD optimizer. Based on the finding, all three give a similiar overal accuracy with the RMSprop being the highest and for the loss section are they all again quite even with the Adam optimizer having the lowest loss. SGD is well known to perform better than Adam with also outperformances on the training data but in our case we had a slight better output with Adam for our dataset and how the model is set up. Lastly, Adam does perform better but if we visualize the loss on the plots compared to SGD, we see that SGD has a even distribution comparing both lines whereas Adam's plots differentiate for each epoch. Lastly, based on the confusion matrix the SGD did better where we would rather have more false positives than false negatives.

3.2. Hypertuning the model loss functions

In this chapter we will use the same steps for setting up the neural network but play around with the hyper paramters to see if we can get a better model by evaluating the loss and accuracy.

We will be following 4 steps:

- 1. Specify some hyper-parameters (the template)
- 2. Train on the training dataset (filling in the parameters)
- 3. Record the validation loss
- 4. Repeat Steps 1 to 3 with a different set of hyper-parameters (many times)

3.2.1 Using CategoricalHinge as loss

Computes the categorical hinge loss between y_true & y_pred.

Neural Network Architecture

Training the Model

```
Epoch 1/100
Epoch 2/100
Epoch 3/100
Epoch 4/100
Epoch 5/100
Epoch 6/100
Epoch 7/100
Epoch 8/100
Epoch 9/100
Epoch 10/100
Epoch 11/100
Epoch 12/100
Epoch 13/100
Epoch 14/100
Epoch 15/100
Epoch 16/100
Epoch 17/100
Epoch 18/100
Epoch 19/100
Epoch 20/100
Epoch 21/100
Epoch 22/100
Epoch 23/100
Epoch 24/100
32/32 [=======================] - 0s 5ms/step - loss: 0.9665 - accuracy: 0.6928 - val loss: 0.9584 - val accuracy: 0.7169
```

```
Epoch 25/100
Epoch 26/100
Epoch 27/100
Epoch 28/100
Epoch 29/100
Epoch 30/100
Epoch 31/100
Epoch 32/100
Epoch 33/100
Epoch 34/100
Epoch 35/100
Epoch 36/100
Epoch 37/100
Epoch 38/100
Epoch 39/100
Epoch 40/100
Epoch 41/100
Epoch 42/100
Epoch 43/100
Epoch 44/100
Epoch 45/100
Epoch 46/100
Epoch 47/100
Epoch 48/100
32/32 [=======================] - 0s 4ms/step - loss: 0.8846 - accuracy: 0.7965 - val loss: 0.8796 - val accuracy: 0.7900
```

```
Epoch 49/100
Epoch 50/100
Epoch 51/100
Epoch 52/100
Epoch 53/100
Epoch 54/100
Epoch 55/100
Epoch 56/100
Epoch 57/100
Epoch 58/100
Epoch 59/100
Epoch 60/100
Epoch 61/100
Epoch 62/100
Epoch 63/100
Epoch 64/100
Epoch 65/100
Epoch 66/100
Epoch 67/100
Epoch 68/100
Epoch 69/100
Epoch 70/100
Epoch 71/100
Epoch 72/100
32/32 [=======================] - 0s 6ms/step - loss: 0.7629 - accuracy: 0.8650 - val_loss: 0.7686 - val accuracy: 0.8356
```

```
Epoch 73/100
Epoch 74/100
Epoch 75/100
Epoch 76/100
Epoch 77/100
Epoch 78/100
Epoch 79/100
Epoch 80/100
Epoch 81/100
Epoch 82/100
Epoch 83/100
Epoch 84/100
Epoch 85/100
Epoch 86/100
Epoch 87/100
Epoch 88/100
Epoch 89/100
Epoch 90/100
Epoch 91/100
Epoch 92/100
Epoch 93/100
Epoch 94/100
Epoch 95/100
Epoch 96/100
32/32 [=======================] - 0s 5ms/step - loss: 0.6964 - accuracy: 0.8738 - val loss: 0.7100 - val accuracy: 0.8539
```

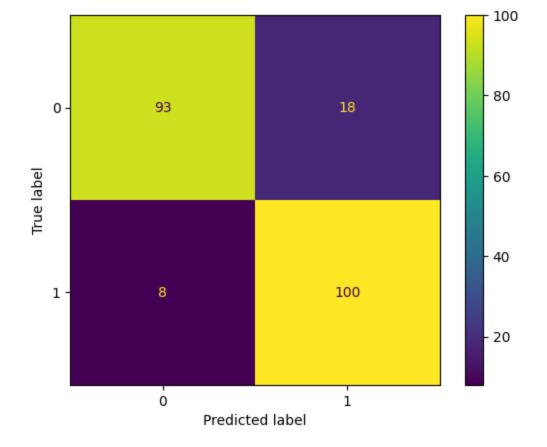
```
Epoch 97/100
Epoch 98/100
Epoch 99/100
32/32 [============== ] - 0s 5ms/step - loss: 0.6916 - accuracy: 0.8738 - val loss: 0.7055 - val accuracy: 0.8539
Epoch 100/100
```

Evaluating the Model

```
In [49]: model_4.evaluate(X_test, Y_test)
    Out[49]: [0.6999451518058777, 0.8812785148620605]
```

Confusion Matrix

```
In [50]: Y pred = model 4.predict(X test).round()
        y pred = np.round(Y pred, 0).tolist()
        confusion matrix(Y test, y pred)
        7/7 [======== ] - 0s 3ms/step
Out[50]: array([[ 93, 18],
               [ 8, 100]], dtype=int64)
In [51]: cm = confusion_matrix(Y_test, y_pred)
        disp = ConfusionMatrixDisplay(confusion_matrix=cm)
        disp.plot()
        plt.show()
```



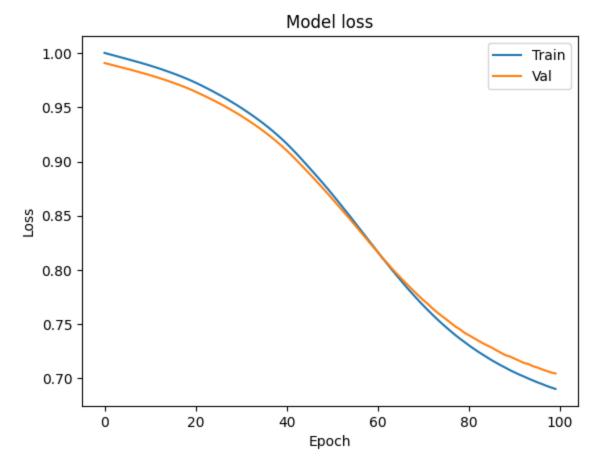
In [52]: print(classification_report(Y_test, y_pred))

	precision	recall	f1-score	support
0	0.92	0.84	0.88	111
1	0.85	0.93	0.88	108
accuracy			0.88	219
macro avg	0.88	0.88	0.88	219
weighted avg	0.88	0.88	0.88	219

Visualizing Loss

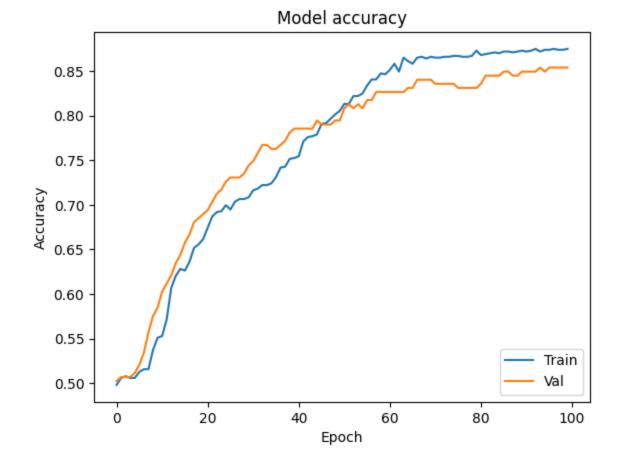
```
In [53]: plt.plot(hist_test_4.history['loss'])
    plt.plot(hist_test_4.history['val_loss'])
    plt.title('Model loss')
    plt.ylabel('Loss')
    plt.xlabel('Epoch')
```

```
plt.legend(['Train', 'Val'], loc='upper right')
plt.show()
```



Visualizing Accuracy

```
In [54]: plt.plot(hist_test_4.history['accuracy'])
    plt.plot(hist_test_4.history['val_accuracy'])
    plt.title('Model accuracy')
    plt.ylabel('Accuracy')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='lower right')
    plt.show()
```



3.2.2 Using poisson as loss

Computes the Poisson loss between y_true and y_pred.

Neural Network Architecture

Training the Model

```
Epoch 1/200
Epoch 2/200
Epoch 3/200
32/32 [===============] - 0s 4ms/step - loss: 0.8435 - accuracy: 0.5812 - val loss: 0.8498 - val accuracy: 0.5753
Epoch 4/200
Epoch 5/200
Epoch 6/200
Epoch 7/200
Epoch 8/200
Epoch 9/200
Epoch 10/200
Epoch 11/200
Epoch 12/200
Epoch 13/200
Epoch 14/200
Epoch 15/200
Epoch 16/200
Epoch 17/200
Epoch 18/200
Epoch 19/200
Epoch 20/200
Epoch 21/200
Epoch 22/200
Epoch 23/200
Epoch 24/200
32/32 [=======================] - 0s 4ms/step - loss: 0.8290 - accuracy: 0.7798 - val loss: 0.8357 - val accuracy: 0.7945
```

```
Epoch 25/200
Epoch 26/200
Epoch 27/200
Epoch 28/200
Epoch 29/200
Epoch 30/200
Epoch 31/200
Epoch 32/200
Epoch 33/200
Epoch 34/200
Epoch 35/200
Epoch 36/200
Epoch 37/200
Epoch 38/200
Epoch 39/200
Epoch 40/200
Epoch 41/200
Epoch 42/200
Epoch 43/200
Epoch 44/200
Epoch 45/200
Epoch 46/200
Epoch 47/200
Epoch 48/200
32/32 [=======================] - 0s 6ms/step - loss: 0.7971 - accuracy: 0.8190 - val loss: 0.8059 - val accuracy: 0.8265
```

```
Epoch 49/200
Epoch 50/200
Epoch 51/200
Epoch 52/200
Epoch 53/200
Epoch 54/200
Epoch 55/200
Epoch 56/200
Epoch 57/200
Epoch 58/200
Epoch 59/200
Epoch 60/200
Epoch 61/200
Epoch 62/200
Epoch 63/200
Epoch 64/200
Epoch 65/200
Epoch 66/200
Epoch 67/200
Epoch 68/200
Epoch 69/200
Epoch 70/200
Epoch 71/200
Epoch 72/200
32/32 [=======================] - 0s 5ms/step - loss: 0.7362 - accuracy: 0.8454 - val loss: 0.7505 - val accuracy: 0.8402
```

```
Epoch 73/200
Epoch 74/200
Epoch 75/200
Epoch 76/200
Epoch 77/200
Epoch 78/200
Epoch 79/200
Epoch 80/200
Epoch 81/200
Epoch 82/200
Epoch 83/200
Epoch 84/200
Epoch 85/200
Epoch 86/200
Epoch 87/200
Epoch 88/200
Epoch 89/200
Epoch 90/200
Epoch 91/200
Epoch 92/200
Epoch 93/200
Epoch 94/200
Epoch 95/200
Epoch 96/200
32/32 [=======================] - 0s 5ms/step - loss: 0.6854 - accuracy: 0.8620 - val_loss: 0.7059 - val accuracy: 0.8265
```

```
Epoch 97/200
Epoch 98/200
Epoch 99/200
Epoch 100/200
Epoch 101/200
Epoch 102/200
Epoch 103/200
Epoch 104/200
Epoch 105/200
Epoch 106/200
Epoch 107/200
Epoch 108/200
Epoch 109/200
Epoch 110/200
Epoch 111/200
Epoch 112/200
Epoch 113/200
Epoch 114/200
Epoch 115/200
Epoch 116/200
Epoch 117/200
Epoch 118/200
Epoch 119/200
Epoch 120/200
32/32 [======================] - 0s 4ms/step - loss: 0.6616 - accuracy: 0.8699 - val loss: 0.6857 - val accuracy: 0.8539
```

```
Epoch 121/200
Epoch 122/200
Epoch 123/200
Epoch 124/200
Epoch 125/200
Epoch 126/200
Epoch 127/200
Epoch 128/200
Epoch 129/200
Epoch 130/200
Epoch 131/200
Epoch 132/200
Epoch 133/200
Epoch 134/200
Epoch 135/200
Epoch 136/200
Epoch 137/200
Epoch 138/200
Epoch 139/200
Epoch 140/200
Epoch 141/200
Epoch 142/200
Epoch 143/200
Epoch 144/200
```

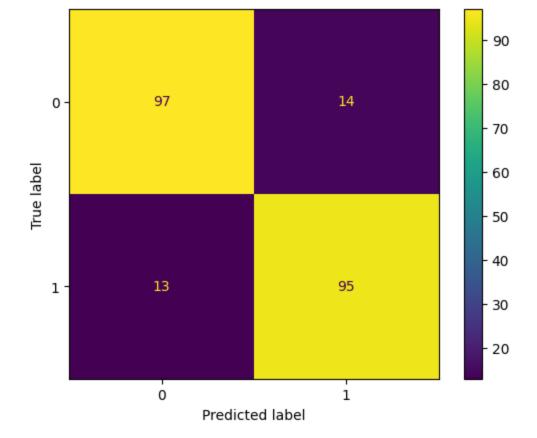
```
Epoch 145/200
Epoch 146/200
Epoch 147/200
Epoch 148/200
Epoch 149/200
Epoch 150/200
Epoch 151/200
Epoch 152/200
Epoch 153/200
Epoch 154/200
Epoch 155/200
Epoch 156/200
Epoch 157/200
Epoch 158/200
Epoch 159/200
Epoch 160/200
Epoch 161/200
Epoch 162/200
Epoch 163/200
Epoch 164/200
Epoch 165/200
Epoch 166/200
Epoch 167/200
Epoch 168/200
32/32 [=======================] - 0s 4ms/step - loss: 0.6417 - accuracy: 0.8777 - val loss: 0.6682 - val accuracy: 0.8539
```

```
Epoch 169/200
Epoch 170/200
Epoch 171/200
Epoch 172/200
Epoch 173/200
Epoch 174/200
Epoch 175/200
Epoch 176/200
Epoch 177/200
Epoch 178/200
Epoch 179/200
Epoch 180/200
Epoch 181/200
Epoch 182/200
Epoch 183/200
Epoch 184/200
Epoch 185/200
Epoch 186/200
Epoch 187/200
Epoch 188/200
Epoch 189/200
Epoch 190/200
Epoch 191/200
Epoch 192/200
32/32 [=======================] - 0s 4ms/step - loss: 0.6358 - accuracy: 0.8845 - val loss: 0.6615 - val accuracy: 0.8584
```

```
Epoch 193/200
Epoch 194/200
Epoch 195/200
Epoch 196/200
Epoch 197/200
Epoch 198/200
Epoch 199/200
Epoch 200/200
```

Evaluating the Model

Confusion Matrix



219

219

219

0.88

0.88

0.88

Visualizing Loss

weighted avg

accuracy macro avg

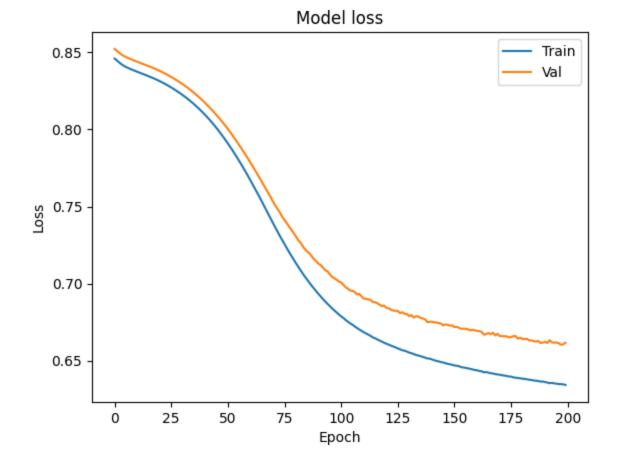
0.88

0.88

0.88

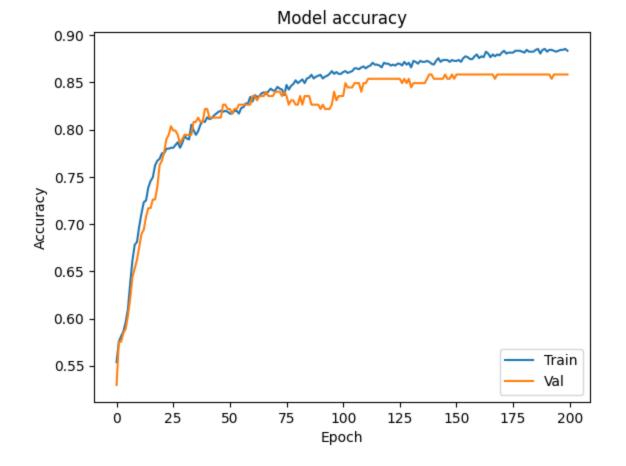
0.88

```
In [62]: plt.plot(hist_test_5.history['loss'])
    plt.plot(hist_test_5.history['val_loss'])
    plt.title('Model loss')
    plt.ylabel('Loss')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='upper right')
    plt.show()
```



Visualizing Accuracy

```
In [63]: plt.plot(hist_test_5.history['accuracy'])
    plt.plot(hist_test_5.history['val_accuracy'])
    plt.title('Model accuracy')
    plt.ylabel('Accuracy')
    plt.xlabel('Epoch')
    plt.legend(['Train', 'Val'], loc='lower right')
    plt.show()
```



3.2.3. Conclusion

We have tried out 2 different loss function to compare to the Binary Crossentropy function which are the poisson and categorical hinge functions. Based on the finding, all three give a similar overal accuracy but comparing the losses which they are, the Binary Crossentropy function performs the best out of the 3 for this data.

3.3. Conclusion on Hypertuning the model

Based on hypertuning the loss function and optimizers we can conclude that the original set hyperparamaters used (Binary Crossentropy & SGD) were the best in comparison to the hyperparameters tried onto the model. The results overall with hypertuning the model came close but were the best ones being the original set hyper parameters which we have concluded based on looking at the loss, accuracy and confusion matrixes.

Reflection

To reflect on this exercise, I have learned how to apply my knowledge and skills in creating a neural network on a random dataset. Although I encountered some difficulties, such as applying the confusion matrix and optimizers, I persisted and eventually gained a better understanding of these concepts.

Additionally, I gained knowledge on new techniques such as Regularization, Early Stopping & Dropout to prevent overfitting of the neural network. Although my neural network in this exercise did not suffer from overfitting, I now have a better understanding of how to address this issue if it were to occur in the future.

Overall, this exercise has provided me with valuable experience and knowledge that I can apply in future projects involving neural networks and machine learning.

Sources

data source that the project is based on: https://www.kaggle.com/c/zillow-prize-1/data

reformed data source: https://drive.google.com/file/d/1GfvKA0qznNVknghV4botnNxyH-KvODOC/view

tutorial/code used to help: https://www.freecodecamp.org/news/how-to-build-your-first-neural-network-to-predict-house-prices-with-keras-f8db83049159/

Helped to set the prediction of the neural network to the right format tot use with also sigmoid included: https://stackoverflow.com/questions/73199505/confusion-matrix-for-binary-classification-with-nn

Steps used for applying the confusion matrix: https://www.jcchouinard.com/confusion-matrix-in-scikit-learn/