**INDIAN INSTITUTE OF TECHNOLOGY PATNA**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING  
  
CS 564: Foundations of Machine Learning**

**Name: Suman Saurabh, Omkar, Atul**

**Roll Number: 2111MC13, 2111MC08, 2111MC05 (in order of name)**

**Course: M Tech. Mathematics and Computing**

**Assignment Documentation**

**First of all we preprocessed the dataset and counter frequency of record for each class as:**

Class

business 510

entertainment 386

politics 417

sport 198

tech 401

**Executed Neural Network model for 5 epochs with relu activation layer function and 50,50 hidden layer nodes in both hidden layer:**

0.2950391644908616

Precsion And Recall value for business Class

Precision: 0.2950391644908616

Recall: 1.0

Precsion And Recall value for tech Class

Precision: 0.2950391644908616

Recall: 0.5978835978835979

Precsion And Recall value for politics Class

Precision: 0.2950391644908616

Recall: 0.41544117647058826

Precsion And Recall value for sport Class

Precision: 0.2950391644908616

Recall: 0.3575949367088608

Precsion And Recall value for entertainment Class

Precision: 0.2950391644908616

Recall: 0.2950391644908616

**We can see the accuracy is very low because model is not converging 5 epochs**

**Next we executed Model for 200 epochs and model is performing extremely well as we can see the output below:**

**Here we are showing Precision and recall Value for each individual class on running model 1 time**

0.9686684073107049

Precsion And Recall value for business Class

Precision: 0.9649122807017544

Recall: 0.9734513274336283

Precsion And Recall value for tech Class

Precision: 0.9685863874345549

Recall: 0.9788359788359788

Precsion And Recall value for politics Class

Precision: 0.9705882352941176

Recall: 0.9705882352941176

Precsion And Recall value for sport Class

Precision: 0.9714285714285714

Recall: 0.9683544303797469

Precsion And Recall value for entertainment Class

Precision: 0.9686684073107049

Recall: 0.9686684073107049

**Then we changed the activation layer function and we can see the model performance has been improved , also overall accuracy value and accuracy of individual class below :**

**Done Validation with 3-Fold Cross Validation:**

**This model is with LBFGS Optimizer and tanh Function in activation layer:**

Performance Measure for Model with Tanh as hidden layer and 50 neurons per Hidden Layer and LBFGS Optimiser

Overall Accuracy for this Model after 3 Fold Validation: 0.9754162094063572

Overall Accuracy value for Business Class: 0.9630742442690128

Overall Accuracy value for Entertainment Class: 0.978973500424193

Overall Accuracy value for Politics Class: 0.9711872089054191

Overall Accuracy value for Sports Class: 0.9951690821256038

Overall Accuracy value for Technical Class: 0.9828101493660005

**Below Model is with ADAM Optimizer and Tanh Activation function in hidden Layer:**

Performance Measure for Model with Tanh as hidden layer and 50 neurons per Hidden Layer and ADAM Optimiser

Overall Accuracy for this Model after 3 Fold Validation: 0.9790751120800382

Overall Accuracy value for Business Class: 0.9730661054597664

Overall Accuracy value for Entertainment Class: 0.9814040955402158

Overall Accuracy value for Politics Class: 0.980887719173896

Overall Accuracy value for Sports Class: 0.9951690821256038

Overall Accuracy value for Technical Class: 0.9754744922630562

**Next Model we tried with SGD Optimizer and Tanh activation function, performance we can see below:**

Performance Measure for Model with Tanh as hidden layer and 50 neurons per Hidden Layer and SGD Optimiser

Overall Accuracy for this Model after 3 Fold Validation: 0.7615258304913478

Overall Accuracy value for Business Class: 0.9920406591344495

Overall Accuracy value for Entertainment Class: 0.7567237624256485

Overall Accuracy value for Politics Class: 0.7520991762462469

Overall Accuracy value for Sports Class: 0.0

Overall Accuracy value for Technical Class: 0.8651609176875135

**Final Model we tried with Relu activation with ADAM Optimiser**

Performance Measure for Model with Relu as hidden layer activation Function and 50 neurons per Hidden Layer and ADAM Optimiser

Overall Accuracy for this Model after 3 Fold Validation: 0.9785551066339244

Overall Accuracy value for Business Class: 0.9730410126162452

Overall Accuracy value for Entertainment Class: 0.9809835121911106

Overall Accuracy value for Politics Class: 0.980887719173896

Overall Accuracy value for Sports Class: 0.9951690821256038

Overall Accuracy value for Technical Class: 0.972551310982162

**CONCLUSION:**

We tried with many combination of Optimizer and activation layer function with two hidden layer and 50 nodes per hidden layer and finally observe following points:

* With 5 Epochs Model is not converging and performing very poorly
* SGD Optimizer is not able to converge with 200 epochs also, it is giving lower overall accuracy of 76 percent.
* **ADAM Optimizer with Tanh as hidden layer is best combination for our dataset with Overall accuracy of 97.9 %**