1 Introduction

- **Group members**
  Vaibhav Anand
  Nikhil Gupta
  Michael Hashe

- **Team name**
  The Breakfast Club

- **Division of labour**

2 Overview

- **Models and techniques tried**
  - **Bullet:** Bullet text.

- **Work timeline**
  - **Bullet:** Bullet text.

3 Approach

- **Data processing and manipulation**
  - **Bullet:** Bullet text.

- **Details of models and techniques**
  - **Bullet:** Bullet text.

4 Model Selection

- **Scoring**
  Due to the binary output classification, we chose to do scoring by classification accuracy (as opposed to loss measure).

- **Validation and Test**
  We used cross-5 validation for all model selection and choosing semi-finalists of the neural-network models, except for choosing the finalists in the neural network models, where we used cross-10 validation. By generating these cross validations, we were able to select the optimal learning parameters, such as regularization terms, early stopping (max iterations), and loss functions. This was done by manually fixing all parameters except for one, optimizing that parameter, then moving on to optimizing another parameter, before circling back to the same parameter. In a way, we performed a manual

step-wise gradient-like descent on parameters. The results of cross validation for various models can be seen in Table # in the Appendix. The parameters variables refer to the API SKlearn's definition of parameters.

5 Conclusion

- **Discoveries**

- **Challenges**

- **Concluding Remarks**

6 Appendix

Table 1: With data normalization, without de-categorization of data (does not include all tests)

Type	K-Folds	Parameters	Classification Accuracy
SVM	5	4-degree polynomial kernel	0.7568
SVM	5	RBF kernel	0.7568
Logistic Regression	5	SAG solver, 25 iterations (converged), 10^{-5} regularization strength	0.774
Logistic Regression	5	SAG solver, 25 iterations (no convergence), $C = 10^0$ regularization strength	0.773
Logistic Regression	5	SAG solver, 100 iterations (no convergence), $C = 10^0$ regularization strength	0.773
Logistic Regression	5	SAG solver, 100 iterations (no convergence), $C = 10^5$ regularization strength	0.773
Logistic Regression	5	SAG solver, 400 iterations (no convergence), $C = 10^0$ regularization strength	0.773
Logistic Regression	5	Liblinear solver, 100 iterations (no convergence), $C = 10^0$ regularization strength	0.774
Ridge Regression	5	$\alpha = 20$ regularization strength, (optimal alpha found by plotting CV vs. alpha)	0.7738
Lasso Regression	5	$\alpha = 10^{-3}$ regularization strength, (optimal alpha found by plotting CV vs. alpha)	0.7718

MLP* Classifier	5	Hidden layers=(200,100), iterations=5 (optimal iterations found by plotting CV vs. iter.)	0.7711
MLP* Classifier	5	Hidden layers=(100), iterations=3 (optimal iterations found by plotting CV vs. iter.)	0.7726
MLP* Classifier	5	Hidden layers=(100, 50, 10), iterations=4 (optimal iterations found by plotting CV vs. iter.)	0.7725
MLP* Classifier	5	Hidden layers=(20), iterations=10 (optimal iterations found by plotting CV vs. iter.)	0.7734
MLP* Classifier	5	Hidden layers=(20, 20), iterations=15 (optimal iterations found by plotting CV vs. iter.)	0.7733
MLP* Classifier	5	Hidden layers=(20, 20,20), iterations=10 (optimal iterations found by plotting CV vs. iter.)	0.7730
MLP* Classifier	5	Hidden layers=(200,100), iterations=5 (optimal iterations found by plotting CV vs. iter.)	0.7711

Table 1: Without data normalization, with de-categorization of data

SVM	5	RBF kernel	0.7707
Logistic Regression	5	Liblinear solver, 50 iterations (converged) $C = 1$ regularization strength	0.7725
Logistic Regression	5	Liblinear solver, 100 iterations (converged) $C = 1$ regularization strength	0.7725
SGD	5	Hinge loss, 1000 iterations, $\alpha = 0.001$ regularization strength	0.7748
SGD	5	Hinge loss, 500 iterations, $\alpha = 0.001$ regularization strength	0.7737
SGD	5	Hinge loss, 100 iterations, $\alpha = 0.001$ regularization strength	0.7727

Table 1: Finalists from 43 cross validations of various layer dimensions and iterations (bounded cross validation by iterations on both sides, such that an increase or decrease in iterations increased validation significantly):

MLP* Classifier	10	Hidden layers=(50, 50), iterations=4 (optimal iterations found by plotting CV vs. iter.)	0.7775

MLP* Classifier	10	Hidden layers=(100, 50, 10), iterations=9 (optimal iterations found by plotting CV vs. iter.)	0.7782
MLP* Classifier	10	Hidden layers=(150, 50), iterations=5 (optimal iterations found by plotting CV vs. iter.)	0.7776
MLP* Classifier	10	Hidden layers=(150, 50, 10), iterations=6 (optimal iterations found by plotting CV vs. iter.)	0.7781

Table 1: *MLP = Multilayered Perceptron. *SGD = Stochastic Gradient Descent. We saw de-categorization improve performance significantly in neural network models, whereas linear models performed slightly worse.