Airline Satisfaction Survey Results

Meg Elliott Fitzgerald and Elise Roberts Scruggs

Data Set - Survey Features

Gender: Gender of the passengers (Female, Male)

Customer Type: The customer type (Loyal customer, disloyal customer)

Type of Travel: Purpose of the flight of the passengers (Personal Travel, Business Travel)

Class: Travel class in the plane of the passengers (Business, Eco, Eco Plus)

Age: The actual age of the passengers

Flight distance: The flight distance of this journey

Departure /Arrival Delay in Minutes: Minutes delayed when departure

Inflight wifi service: Satisfaction level of the inflight wifi service (0:Not Applicable;1-5)

Departure/Arrival time convenient: Satisfaction level of Departure/Arrival time convenient (0:Not Applicable;1-5)

Ease of Online booking: Satisfaction level of online booking (0:Not Applicable;1-5)

Gate location: Satisfaction level of Gate location (0:Not Applicable;1-5)

Food and drink: Satisfaction level of Food and drink (0:Not Applicable;1-5)

Online boarding: Satisfaction level of online boarding(0:Not Applicable;1-5)

Data Set - Survey Features Continued

Seat comfort: Satisfaction level of Seat comfort (0:Not Applicable;1-5)

On-board service: Satisfaction level of On-board service (0:Not Applicable;1-5)

Leg room service: Satisfaction level of Leg room service (0:Not Applicable;1-5)

Baggage handling: Satisfaction level of baggage handling (0:Not Applicable;1-5)

Check-in service: Satisfaction level of Check-in service (0:Not Applicable;1-5)

Inflight service: Satisfaction level of inflight service (0:Not Applicable;1-5)

Cleanliness: Satisfaction level of Cleanliness (0:Not Applicable;1-5)

Inflight entertainment: Satisfaction level of inflight entertainment (0:Not Applicable;1-5)

TARGET: Satisfaction: Airline satisfaction level(Satisfaction, neutral or dissatisfaction)

Likert Scale Features

- Named after American psychologist Rensis Likert in 1932
  - Scale of agreement (typically 1 to 5)
- Likert-type scales defined by Paul Lavrakas in 2008
  - Allows respondents to indicate degree of satisfaction (typically 1 to 5)

How satisfied are you with the security control at the airport?

Very dissatisfied	Somewhat dissatisfied	Neither dissatisfied nor satisfied	Somewhat satisfied	Very satisfied
○	○	○	○	○
1	2	3	4	5

Using Likert-type scales for Classification

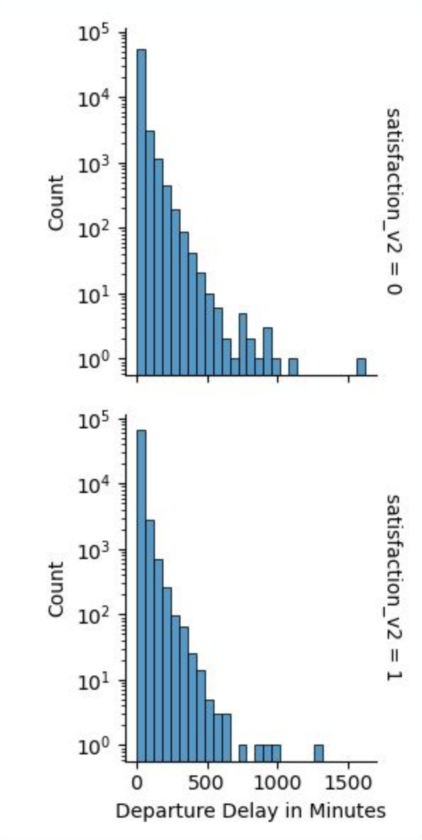
- Endresen and Janda (2016) discovered that Random Forest classification models were the best classical models to use for Likert or Likert-type scale features
- Will a Multilayer Perceptron Network perform better than a Random Forest?

Exploratory Data Analysis - Categorical

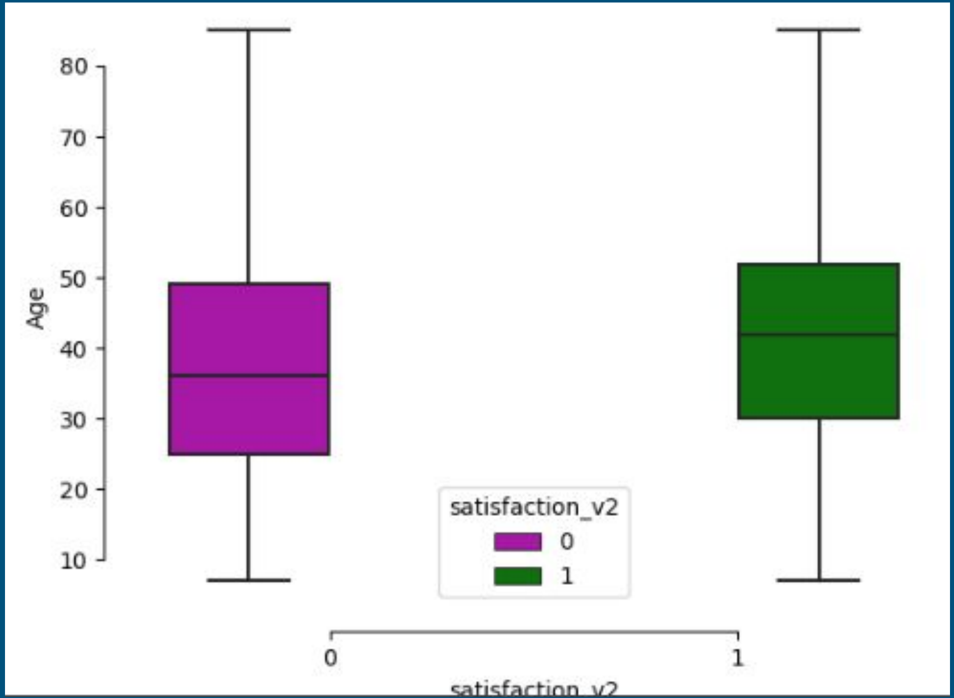
- Gender
  - Women more satisfied with airline than men
- Customer Type
  - Loyal Customers more satisfied than Disloyal Customers
- Type of Travel
  - Business travel passengers more satisfied than Personal travel passengers
- Travel Class
  - Business Class and Economy Plus passengers more satisfied than Economy Class passengers

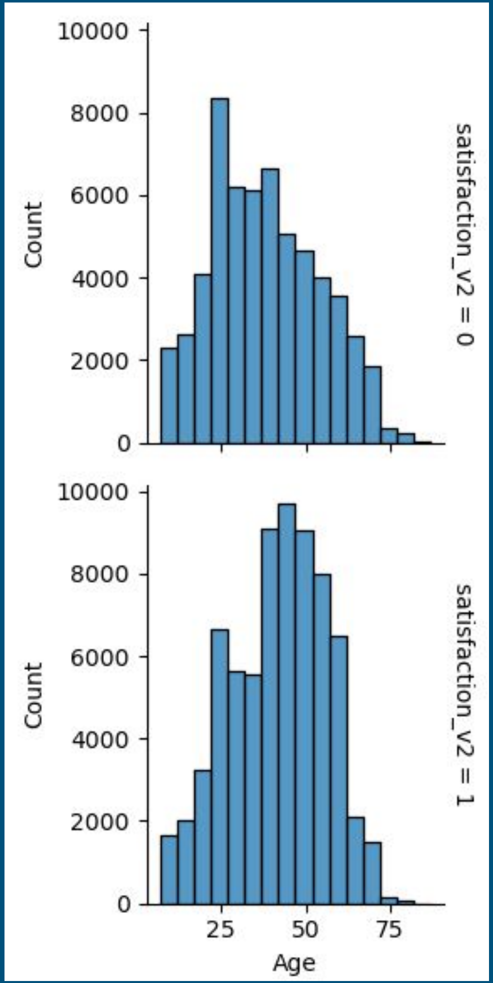
Exploratory Data Analysis - Continuous

- Departure and Arrival Delays
  - Skewed Distributions with no significant correlations, even with modifications to a log scale
  - Therefore, we excluded this feature from the final model
- Flight Distance
  - Not highly correlated with satisfaction
- Age
  - Satisfaction not highly correlated with Age



Investigation of Age Feature





Likert-type - Highest Correlations

	Seat comfort	satisfaction_v2
0	0	0.997908
1	1	0.450915
2	2	0.357794
3	3	0.356131
4	4	0.651845
5	5	0.992064

	Inflight entertainment	satisfaction_v2
0	0	0.660714
1	1	0.210656
2	2	0.170363
3	3	0.199188
4	4	0.719870
5	5	0.952064

Likert-type - Moderate/Average Correlations

	Gate location	satisfaction_v2
0	0	1.000000
1	1	0.610926
2	2	0.580705
3	3	0.463065
4	4	0.497850
5	5	0.655479

	Food and drink	satisfaction_v2
0	0	0.779635
1	1	0.508473
2	2	0.432713
3	3	0.428612
4	4	0.590254
5	5	0.780084

	Inflight wifi service	satisfaction_v2
0	0	0.446154
1	1	0.268507
2	2	0.502356
3	3	0.509557
4	4	0.638368
5	5	0.669114

	Online support	satisfaction_v2
0	0	0.000000
1	1	0.295464
2	2	0.296755
3	3	0.282690
4	4	0.680481
5	5	0.773152

	Ease of Online booking	satisfaction_v2
0	0	0.000000
1	1	0.192730
2	2	0.286368
3	3	0.357053
4	4	0.717813
5	5	0.765499

	On-board service	satisfaction_v2
0	0	0.000000
1	1	0.265825
2	2	0.340831
3	3	0.410290
4	4	0.647394
5	5	0.765692

Likert-type - Moderate/Average Correlations

	Leg room service	satisfaction_v2
0	0	0.692308
1	1	0.283384
2	2	0.376378
3	3	0.371746
4	4	0.673345
5	5	0.708523

	Departure/Arrival time convenient	satisfaction_v2
0	0	0.542444
1	1	0.586154
2	2	0.540268
3	3	0.539680
4	4	0.524607
5	5	0.556450

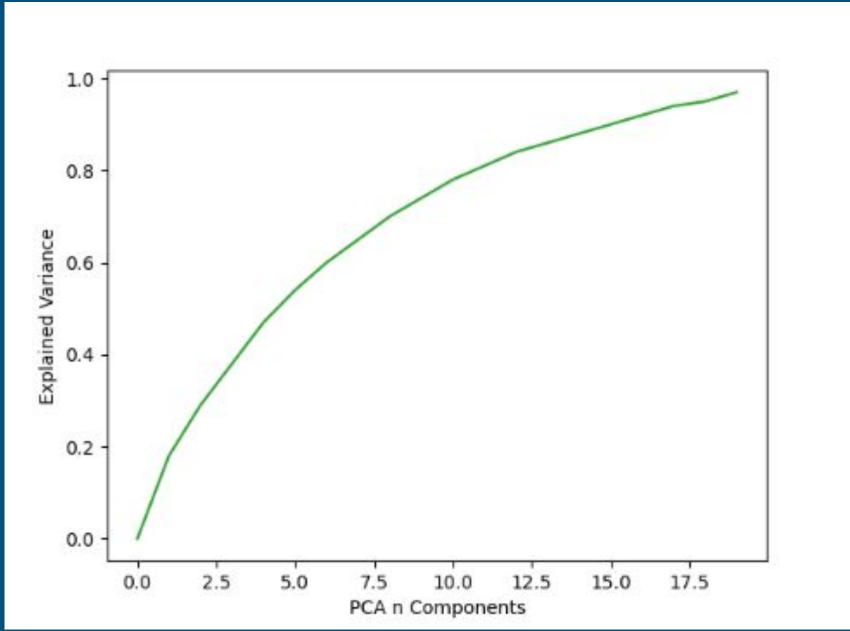
	Cleanliness	satisfaction_v2
0	0	0.000000
1	1	0.403047
2	2	0.404760
3	3	0.317899
4	4	0.586767
5	5	0.731698

	Online boarding	satisfaction_v2
0	0	0.000000
1	1	0.264925
2	2	0.281309
3	3	0.549557
4	4	0.652527
5	5	0.731715

	Baggage handling	satisfaction_v2
0	1	0.422574
1	2	0.396250
2	3	0.315119
3	4	0.588480
4	5	0.735817

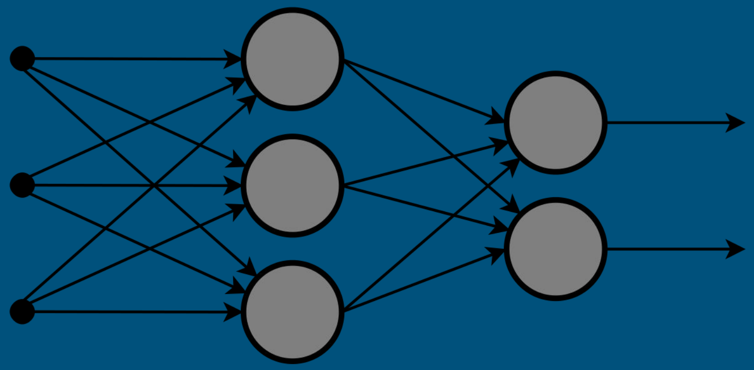
	Checkin service	satisfaction_v2
0	0	0.000000
1	1	0.316277
2	2	0.332643
3	3	0.567485
4	4	0.577010
5	5	0.735763

Principal Component Analysis



Multilayer Perceptron Neural Network

Binary Classification Network



3 Layer MLP - Adam

- First-order, gradient-based optimization of stochastic objective functions
  - Straightforward to implement
  - Computationally efficient
  - Little memory requirements

	precision	recall	f1-score	support
0	0.91	0.91	0.91	11692
1	0.92	0.92	0.92	14206
accuracy			0.92	25898
macro avg	0.91	0.92	0.92	25898
weighted avg	0.92	0.92	0.92	25898

3 Layer MLP - SGD

- Stochastic Gradient Descent Backpropagation

	precision	recall	f1-score	support
0	0.90	0.91	0.90	11692
1	0.92	0.92	0.92	14206
accuracy			0.91	25898
macro avg	0.91	0.91	0.91	25898
weighted avg	0.91	0.91	0.91	25898

3 Layer MLP - L-BFGS

- Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS)
    - Quasi-Newtonian Method
    - Uses the approximated second order gradient information
- Best model with 15 neurons in each layer

	precision	recall	f1-score	support
0	0.91	0.93	0.92	11692
1	0.94	0.93	0.93	14206
accuracy			0.93	25898
macro avg	0.93	0.93	0.93	25898
weighted avg	0.93	0.93	0.93	25898

Modifications of 3 Layer LBFGS Model

- Increase the neurons in the second layer only
  - Number of neurons in each layer: 10 - 15 - 10
- Increase the neurons in all layers
  - Number of neurons in each layer: 15 - 15 - 15
  - Number of neurons in each layer: 20 - 20 - 20
- Change activation function
  - Logistic regression
- Using SGD, change momentum and alpha
  - Momentum 0.95 and alpha 0.01
  - Momentum 0.95 and alpha 0.01 with logistic regression

4 Layer MLP

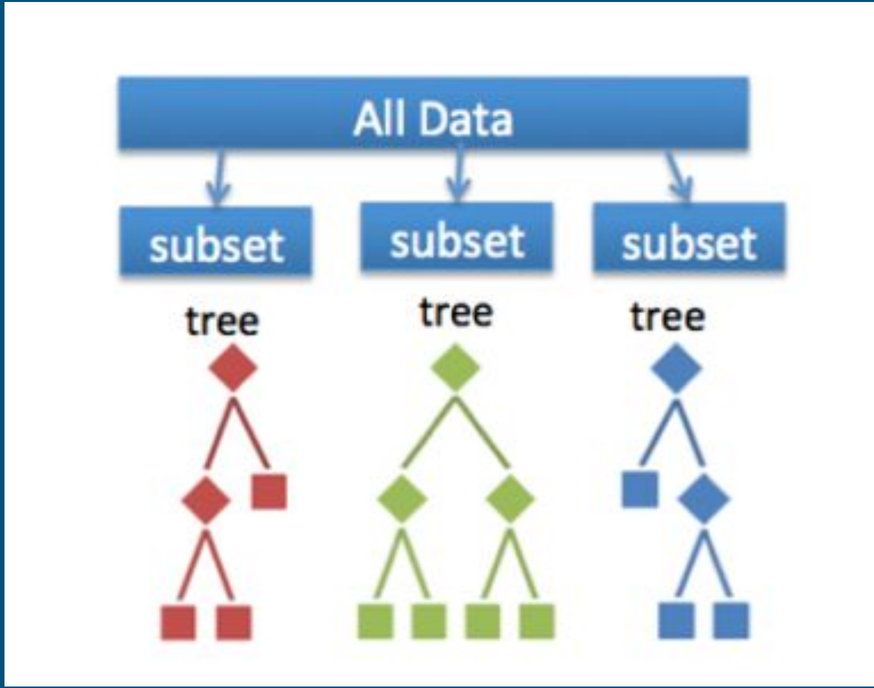
- Using L-BFGS
  - Number of neurons in each layer: 10 - 10 - 10 - 10 as well as 15 - 15 - 15 - 15

	precision	recall	f1-score	support
0	0.90	0.91	0.91	11692
1	0.93	0.92	0.92	14206
accuracy			0.92	25898
macro avg	0.92	0.92	0.92	25898
weighted avg	0.92	0.92	0.92	25898

Classical Model

Random Forest Classifier

- Accuracy of 0.92



Results

- Best Model - 3 Multilayer Perceptron Network with L-BFGS optimization method and layer sizes of 15 - 15 - 15

- Recommend using Random Forest for Likert-type scale features

Conclusion

- Multilayer Perceptron Network can be used to classify Likert-type scale features from survey responses
- Agree with Endreson and Janda - recommend using classical Random Forest classification model

References

D, J. (2018, June 10). *Passenger Satisfaction*. Kaggle. Retrieved on 13 June 2021 from https://www.kaggle.com/johndddddd/customer-satisfaction.

Das, S. (2021, June 11). *Airline Passenger Satisfaction - Prediction > 95%*. Kaggle. Retrieved on 13 June 2021 from https://www.kaggle.com/codesagnik/airline-passenger-satisfaction-prediction-95.

Endresen, A., & Janda, L. A. (2016). Five statistical models for Likert-type experimental data on acceptability judgments. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 3(2). Retrieved on 20 June 2021 from https://doi.org/10.1558/jrds.v2i1.2269

Goodfellow, I., Bengio, Y., & Courville, A. (2017). 6.5 Back-Propagation and Other Differentiation Algorithms. In *Deep Learning* (pp. 200–220). MIT Press.

Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (2016). *Neural network design*. s. n. https://hagan.okstate.edu/nnd.html.

Jais, I. K. M., Ismail, A. R., & Nisa, S. Q. (2019). Adam Optimization Algorithm for Wide and Deep Neural Network . *Knowledge Engineering and Data Science (KEDS)* , 2(1), 41–46. https://core.ac.uk/download/pdf/287322851.pdf

Kingma, D. P., & Ba, J. (2017, January 30). *Adam: A Method for Stochastic Optimization.* arXiv.org. https://arxiv.org/abs/1412.6980.

Lavrakas, P. J. (2008). *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: SAGE Publications. https://doi.org/10.4135/9781412963947

References Continued

Likert, R. (1932). A Technique for the Measurement of Attitudes. Doctoral dissertation. Columbia University. Series Archives of Psychology 22: 5–55. *NY: The Science Press.* Retrieved on 20 June 2021 from http://www.voteview.com/pdf/Likert_1932.pdf

Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Mathematical Programming: Series A and B*, 45(1-3), 503–528. https://doi.org/10.5555/3112655.3112866

Moritz, P., Nishihara, R., & Jordan, M. I. (2016, April 13). *A Linearly-Convergent Stochastic L-BFGS Algorithm*. arXiv.org. https://arxiv.org//abs/1508.02087.

Najafabadi, M. M., Khoshgoftaar, T. M., Villanustre, F., & Holt, J. (2017). Large-scale distributed L-BFGS. *Journal of Big Data,* 4(22). https://doi.org/10.1186/s40537-017-0084-5

Rafati, J., & Marcia, R. F. (2018). Improving L-BFGS Initialization for Trust-Region Methods in Deep Learning. *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. https://doi.org/10.1109/icmla.2018.00081

Rumelhart, D. E., Group, P. D. P. R., & McClelland, J. L. (1986). *Parallel Distributed Processing, Explorations in the Microstructure of Cognition: Foundations (Vol. 1)*. MIT Press.

Saad, D. (1998). *On-line learning in neural networks*. Cambridge.

Shamir, O. & Zhang,T. (2013). Stochastic Gradient Descent for Non-smooth Optimization: Convergence Results and Optimal Averaging Schemes. *Journal of Machine Learning Research: W & CP*: 28, p. 71-79. http://proceedings.mlr.press/v28/shamir13.pdf.

Xiao, Y., Wei, Z., Wang, Z. (2008). A limited memory BFGS-type method for large-scale unconstrained optimization. *Computers & Mathematics with Applications*, 56(4), 1001-1009. https://doi.org/10.1016/j.camwa.2008.01.028.