### To what extent can the Naïve Bayes machine learning algorithm be used to help filter email spam ?

Candidate Name: AYAAN MECKLAI

Centre Number: IN693

Candidate Number: 0006

Syllabus Code: 9980

Word Count: 4405

**BIBLIOGRAPHY**

Aladdinpersson. “Aladdinpersson/Machine-Learning-Collection: A Resource for Learning about ML, DL, Pytorch and Tensorflow. *GitHub*, https://github.com/aladdinpersson/Machine-Learning-Collection.

Aladdinpersson. “Machine-Learning-Collection/Naivebayes.py at Master · Aladdinpersson/Machine-Learning-Collection.” *GitHub*, https://github.com/aladdinpersson/Machine-Learning-Collection/blob/master/ML/algorithms/naivebayes/naivebayes.py.

Brownlee, Jason. “How to Choose a Feature Selection Method for Machine Learning.” *Machine Learning Mastery*, 20 Aug. 2020, https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/.

Dada, Emmanuel Gbenga, et al. “Machine Learning for Email Spam Filtering: Review, Approaches and Open Research Problems.” *Heliyon*, Elsevier, 10 June 2019, https://www.sciencedirect.com/science/article/pii/S2405844018353404.

“Data Preprocessing in Data Mining.” *GeeksforGeeks*, 29 June 2021, https://www.geeksforgeeks.org/data-preprocessing-in-data-mining/.

*Email Spam Filtering Using Supervised Machine Learning Techniques*. https://www.researchgate.net/publication/50235326\_Email\_Spam\_Filtering\_using\_Supervised\_Machine\_Learning\_Techniques.

*Email Spam Filtering Using Supervised Machine Learning Techniques*. https://www.researchgate.net/publication/50235326\_Email\_Spam\_Filtering\_using\_Supervised\_Machine\_Learning\_Techniques.

Hrouda-Rasmussen, Stefan. “(Gaussian) Naive Bayes.” *Medium*, Towards Data Science, 7 May 2021, https://towardsdatascience.com/gaussian-naive-bayes-4d2895d139a.

Joby, Amal. “What Is Training Data? How It's Used in Machine Learning.” *Learn Hub*, https://learn.g2.com/training-data.

Makispl. “MAKISPL/SMS-SPAM-Filter-Naive-Bayes: A Naive Bayes Spam Filter Developed in Python.” *GitHub*, https://github.com/makispl/SMS-Spam-Filter-Naive-Bayes.

“Scoring Data.” *DataRobot AI Cloud*, 20 July 2022, https://www.datarobot.com/wiki/scoring/.

“Scoring  |  Machine Learning  |  Google Developers.” *Google*, Google, https://developers.google.com/machine-learning/recommendation/dnn/scoring.

*SPAM Message Classification Based on the Naïve Bayes Classification ...* http://www.iaeng.org/IJCS/issues\_v46/issue\_1/IJCS\_46\_1\_05.pdf.

Stedman, Craig, and Adam Hughes. “What Is Data Mining?” *SearchBusinessAnalytics*, TechTarget, 7 Sept. 2021, https://www.techtarget.com/searchbusinessanalytics/definition/data-mining.

Techopedia. “What Is Tokenization? - Definition from Techopedia.” *Techopedia.com*, Techopedia, 3 June 2021, https://www.techopedia.com/definition/13698/tokenization.

Techopedia. “What Is Training Data? - Definition from Techopedia.” *Techopedia.com*, Techopedia, 17 Feb. 2022, https://www.techopedia.com/definition/33181/training-data#:~:text=Training%20data%20is%20an%20extremely,the%20training%20data%20is%20labeled.

Techopedia. “What Is Training Data? - Definition from Techopedia.” *Techopedia.com*, Techopedia, 17 Feb. 2022, https://www.techopedia.com/definition/33181/training-data.

“What Is Exploratory Data Analysis ?” *GeeksforGeeks*, 26 Apr. 2022, https://www.geeksforgeeks.org/what-is-exploratory-data-analysis/.