125   COVID-19 PREDICTION USING SYMPTOMS

Abstract

The recent COVID-19 outbreak, which was triggered by the new coronavirus SARS-Cov2, is a pressing global concern. It would be useful to determine which symptoms presented by suspected infection patients are the best predictors of a positive diagnosis to assist in making smart decisions on treatment needs.

Introduction

Develop prediction models that combine several features to estimate the risk of infection and assist medical staff worldwide in triaging patients, especially in the context of limited healthcare resources.

Research Question(s)

- How big was the sample
- Choosing the data set
- Which decision tree to use
- Link between predicting covid-19 based on gender

Materials and Methods

K Nearest Neighbor

Naive Bayes Algorithm

DECISION TREES

Results

ML Algorithm	Accuracy	Precision	Recall	F1-Score	Confusion Matrix
KNN	91.88	0 − 0.96 1 − 0.51	0 − 0.96 1 − 0.52	0 − 0.96 1 − 0.52	[25992 1220] [1190 1295]
Decision Trees	95.1	0 − 0.96 1 − 0.77	0 − 0.98 1 − 0.60	0 − 0.97 1 − 0.68	[26774 438] [994 1491]
Naive Bayes	91.88	0 − 0.96 1 − 0.51	0 − 0.96 1 − 0.52	0 − 0.96 1 − 0.52	[25992 1220] [1190 1295]

Resources:
- https://www.kdnuggets.com/2020/05/model-evaluation-metrics-machine-learning.html
- https://www.nature.com/articles/s41746-020-00372-6
- Introduction to machine learning (4 th edition) by Ethem Alpaydin, The MIT Press, 2020. ISBN: 978- 0262043793
- https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-021-06810-4

Internship Poster should include the following components.

Company – Kennesaw State University

Introduction– Students pursuing Masters in Kennesaw State University

Highlights

- Python 3 & AWS
- The final code is not any specific IDE dependent. We used our preferred IDE such as Jupyter Notebook, Visual Studio Code, etc. environment in the production phase.
- Used packages such as:
- NumPy
- Pandas
- Datetime
- OS
- Sklearn
- Matplotlib

Experience

Pursuing a career in AI is our ambition. So, getting exposure to this project helped us learn many ml algorithms and determine their accuracies.

Acknowledgments

Dr. Mahmut Karakaya

Future Career Plans

We would like to further extend our career in AI by doing research in this area and pursue phD.

Conclusions

Decision trees proved to be the best in predicting covid-19 based on symptoms.

Acknowledgments

Dr. Mahmut Karakaya

Contact Information

ygurjala@students.kennesaw..edu
spethe@students.kennesaw.edu
cveerave@students.kennesaw.edu

References

- https://www.kdnuggets.com/2020/05/model-evaluation-metrics-machine-learning.html
- https://www.nature.com/articles/s41746-020-00372-6
- Introduction to machine learning (4 th edition) by Ethem Alpaydin, The MIT Press, 2020. ISBN: 978-0262043793
- https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-021-06810-4

KENNESAW STATE UNIVERSITY
COLLEGE OF COMPUTING AND SOFTWARE ENGINEERING

Author(s)- yagna gurjala, chalamaya batchu, shruthi pethe
Advisors(s)- Dr. Mahmut Karakaya