COVID-19 Prediction using machine learning

Jay Patel Jithu Manoj

Lakehead University Lakehead University

[Jpatel22@lakeheadu.ca](mailto:Jpatel22@lakeheadu.ca) [jmanoj@lakeheadu.ca](mailto:jmanoj@lakeheadu.ca)

**Abstract**

Logistic regression(LR) continues to be one of the most widely used methods in data mining in general and binary data classiﬁcation in particular. This paper is focused on providing an overview of the most important aspects of Logistic regression when used in machine learning, speciﬁcally from an algorithmic perspective and how logistic regression can be applied to imbalanced and rare events data.

**1.Keywords**

Logistic regression, data mining, classification, imbalanced data

**2.Introduction**

Logistic Regression (LR) is one of the most important statistical and data mining techniques employed by statisticians and researchers for the analysis and classiﬁcation of binary and proportional response data sets . Some of the main advantages of LR are that it can naturally provide probabilities and extend to multi-class classiﬁcation problems . Another advantage is that most of the methods used in LR model analysis follow the same principles used in linear regression . What’s more, most of the unconstrained optimization techniques can be applied to LR.

Here the problem accepts several inputs from the user regarding the symptoms of covid-19 and the system analyse the inputs and predict the result. This paper provides an overview of some of the algorithms and the corrections that enable LR to be both fast and accurate from a machine learning point of view. It is by no means an exhaustive survey of all the LR techniques in data mining.

**3.Website development**

After developing the model we started to develop the websites. To accomplish this, we created a templates directory in our project containing a base.html file that 4 other pages would inherit from. A website is dedicated to show the information about or team members and job allocations.

We also used a method called Bootstraps to create the header of the website that allows us to switch between all the other pages easily. The outline page of the website is used to describe the model and methods used to accomplish its implementation. Interface page allows the user to enter the data regarding their symptoms into the webpage. From there we used the request method to read the data and search the train/test file to accurately show the predicted COVID result.

**4.Docker file development and deploying GCP**

Firstly I created a docker file then I pushed that to my google cloud console after creating a VM instance. After that I installed the docker certificates on the cloud shell. Then you can access the docker from remote repo, that you can build it and run docker. Then you can access an ip address which your app will be running.

**5. Deploying Kubernetes**

After running the app on GCP we created a database in mysql and then deployed it on Kumbernetes. It is used to store data that we get from interface.

**6.References**

1. <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/forecasting.html#:~:text=Why%20COVID%2D19%20Forecasting%20Is%20Important&text=Forecasts%20of%20deaths%2C%20hospitalizations%2C%20and,19%20pandemic%20in%20coming%20weeks>
2. [https://www.zdnet.com/article/moving-to-iaas-an-overview/ - :~:text=In%20a%20typical%20IaaS%20deployment,'spun%20up'%20on%20demand%2C](https://www.zdnet.com/article/moving-to-iaas-an-overview/#:~:text=In%20a%20typical%20IaaS%20deployment,'spun%20up'%20on%20demand%2C)
3. <https://www.forbes.com/sites/janakirammsv/2019/04/16/why-you-should-consider-google-ai-platform-for-your-machine-learning-projects/?sh=526559cf22ab>