Search Parasparam 2020

Parasparam 2020

[View All Results ()](#)

V

Paraspai

HOME
FAQS
MESSAGES
NEWS

Subscribe

3

POINTS

AI based multi-factor risk assessment expert system for admitting HPE employees during COVID-19

(D868)

👤 [**Team - Yash Dalvi**](#)   Submitted: Jun 4 2020   Category: [Big Data and Analytics](#)   Status: Pending

artificial intelligence   disaster management strategy   framework

☐ ☐ ☐ ☐

Clear Vote

Find Teammates

| **Submission** | **+** |
| **Tea ı** | **Add/Edit** |

Y   C   S   V

Details

Points:	3
Votes:	3
Rank:	3
Unique Views:	6
Total Views:	57
Comments:	0
Favorited:	0

Linked Submissions:
No Linked Submissions

Request To Join Team

Description	Attachments (2)	Votes

The world is facing an unprecedented crisis with the deadly spread of COVID-19 virus. As of May 2020, the total number of people found positive with COVID are 5.23M and the deaths reported are about 338K. The virus has hit the global economy which has resulted in stagnation of manufacturing, production, travel and global supply chain industries.

As a result, a lot of companies are finding it really difficult to operate in such drastic conditions. The major challenge the industries are facing is the strategy to open the offices for their employees considering the risk of spreading the Corona virus in the company. In addition, it is a challenge for the manager to select few employees to allow work in the office as Government has asked to organization to operate with limited percentage (30%) of their strength. The proposed solution directly addresses this problem and helps organization to assess the risk of COVID-19 if it allows particular employee to enter the campus. The solution captures all the relevant data such as employee's immunity, current medical condition/symptoms, past medical data, travel history and the commute route to the office. The solution also assess the risk involving the inter-state travels and future travel of the employee. So considering the urgency of the matter an Artificial Intelligence(AI) based solution is proposed to help the mangers to take a data-driven and informed decision of selecting the employees that can be admitted to the office premises.

The similar approach is planned to be followed by site management and leader ship team @HPE and the proposal mentioned in this paper will certainly make decision making process much more efficient and easier.

Expertise Required

| Artificial Intelligence | Data Science |

What problem are you addressing with your project?

The COVID19 crisis has already crippled the world economies. It is now been predicted that the situation caused by this pandemic would be worse than the 2008 recession. In such tough times, all the world is busy battling and preparing strategies through which could useful for a sustainable recovery with minimum risk involved. Many of the economic industries and corporate offices are trying to devise plans which would help them to deal with the crisis. Keeping this in mind, we thought that one of most challenging problem that HPE would be facing is to access the overall impact and risk factors for its employees. As we all are trying to recover from the pandemic in small and different phases, the campus premises will be allowed to accommodate more number of employees as we go forward. This decision

of whether a given employee must be admitted to the campus needs to be taken with utmost amount of informed data and by accessing all the possible risk factors associated with the employee.

With state governments taking stricter measures with home and institutional quarantine in place, the main challenge would be to predict and evaluate the risks associated with future inter-state travels.

Along with the medical and immunity risk associated, we are trying to measure all the possible risk factors that can occur during the travel, so that we reduce the overall risk and that HPE employees in other states can commute back in the most favorable and secured conditions for them.

What is the solution you are proposing

Our solution tries to take in account all the possible ways in which we can measure the overall risk involved for a given employee. We have prepared a draft employee survey to get required information for further analysis.

Survey Link
: https://docs.google.com/forms/d/e/1FAIpQLScEMxmGhERS2GmH1Y_n94ggwg9dYwiisEy63XvtIyn-Eky0SQ/viewform (Attached sample survey in attachments section)

We have categorized risk assessment in three parts -
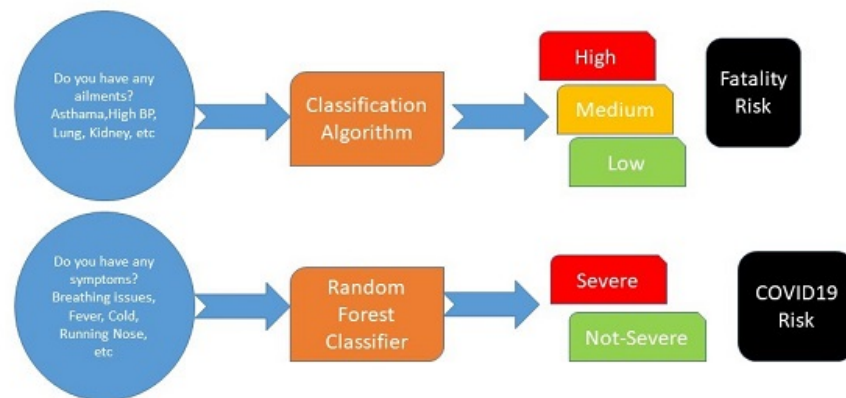
1. Risks Based on Medical History of an Employee -

By surveying various scientific research articles related to COVID19, we have gathered information related to various medical aliments like asthama, respiratory diseases, heart problems, etc which can aggravated severely by COVID19. Our modeled solution would contain a well-formed questionnaire, which the mangers can hand their employees. Once all the data points are received from the questionnaire, our solution would categorize the health risk as low, medium or severe for a given employee. The aim here is that we should not admit any employee who has already these identified fatal aliments, as any changes of exposing to COVID19 can severely put the employees life at high risk.

2. Risk based on presence of existing COVID19 symptoms -

As in point number 1, we have prepared a similar questionnaire, which for a given employee, helps us to examine whether the employee is already exposed to COVID19. This questionnaire mainly include the COVID19 symptoms like presence of dry-cough, breathing issues, fever, etc. This is very similar to the aarogya setu app, the solution gives result as the risk involved in exposure to COVID19 is sever or mild. The final outcome is achieved by training grid search over random forest classifier and xgBoost classifier. A published dataset on kaggle was used as the training data. The statistics for the model were -

Accuracy   0.7509 $ACC = (TP + TN) / (P + N)$
F1 Score   0.8577 $F1 = 2TP / (2TP + FP + FN)$
Confusion Matrix generated from cleaned data and developed model:
array([[71363,   0],
     [23677,   0]])
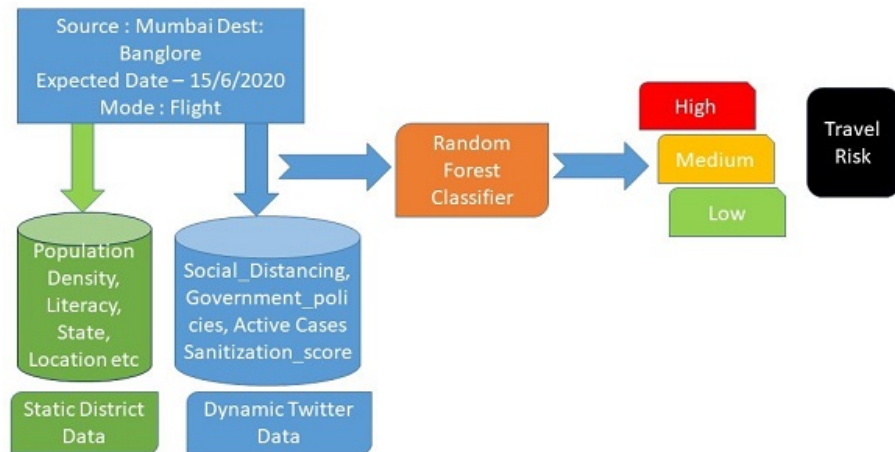
3. Risk predicted for travelling for a given date in future -

This is the most important part of the problem which our solution tries to solve. For a given employee, supposedly, he has travel back to his native, and now needs to travel back to Bangalore campus. The journey date will approximately after 2-3 weeks but the the risk evaluation needs to be right now, for the manager. We have tried to predict the risk using Artificial Intelligence. The core part of this solution lies in the construction of its dataset. We have build an extensive dataset which, for every district present in India, contains the following information - District population density, literacy rate, population growth rate, state, sex ratio, active COVID19 cases, confirmed corona cases, recovered cases, delta for active, confirmed and recovered cases. The data for the features such as population density, literacy, state etc is collect from the official government census. The data related to COVID19 cases is fetched through available API's use web-scrapping and gather the information from the official government site.

Considering the dynamic nature COVID19 cases that are changing and that too at a rapid rate, it is a real challenge to measure this change rate, co-relate different events happening in real time and find insights that can help us in understanding its nature and its trend. To collect, measure and analyze this data, we have resolved to twitter. Twitter has tremendous amount of unstructured data posted through different organizations like news channels, NGO's etc government entities like Railways, Medical Council etc as well as the common public. We believe that if we are able to co-related timely events, we will definitely be able to make sense out of it. For example, if we consider a tweet like - 3 passengers who were detected to be COVID19 positive in Bangalore traveled from Mumbai via flight. We can conclude lot of useful information from the tweet like the date of travel, source and the destination cities involved - Bangalore,Mumbai, mode of transport -flight, etc Using such type of data we plan to prepare a feature model which can keep score patients traveled for particular cities and their mode of transport. This data also helps us to measure other things like social distancing, cleanliness, sanitization norms that are being followed at airports and their success rate.Using such data, if we find that if there are certain airports, cities which are constantly coming under this radar, we can flag that travel route to be unsafe/risky.

Another feature which we are trying to build is the general notion for COVID19 cases, whether a particular district is witnessing increase or decrease in the number of cases by capturing the sentiment tweets about COVID19. A huge number of tweets for a particular district containing words like increasing, problematic, more etc. can convey sense that the cases are increasing while if the tweets contain "happy" words like recovered, well, discharged, etc can convey the trend that the place is witnessing decrease in the number of cases. We are also trying to include government decisions like type of lockdown, the relaxation offered, etc to measure the factors which really contribute to the outcome of the COVID19 cases.

What is the evidence that solution works

For each of three solutions -

1. We have formulated a classification algorithm to classify the disease into low, medium and high risk category.

2. We have trained Random Forest Algorithm with xgboost with an accuracy of 75%

3. We have trained Random Forest Regression Algorithm with an accuracy of 40% without including twitter data. This accuracy score is expected to increase much further with the inclusion of twitter data

What are the other competitive approaches ?

Lot of the technologist are trying to solve the problem using AI. Most of them are focused on measuring and predicting the growth of COVID19 in particular area. However, we did not find any solution that measures the risk associated with the travel between two locations.

What is the current Status

Work in Progress

What are the next steps

Our next step involve in enriching our dataset for better accuracy. We plan to take in other factors like longitude, latitude of the districts, nearest metropolitan city distance, temperature, medical facilities available, presence of airport, strength of railway station, daily travel frequency, etc

For the features which depend on the twitter data, we can use the images as well. Using deep learning models, we can extract information whether social distancing was followed or not, how much of trading is allowed in the city markets , people movement, etc

We also propose that such type of data collection strategy can be benefited to any type of calamity faced in the future as the overall architecture of proposed solution can be used as a generic framework to implement a disaster management strategy.

We(Most of us) have got instructions on the activities and process to be followed for site visit from site management team and leadership team. E.g portal for employee survey, Access date-time etc. and thus we are seeing potential use of overall risk tracking mechanism mentioned in this paper since we will be having complete data available in system itself and the analysis performed on the employee data will certainly help in getting more insights and efficient decision making for management and security authorities.

Acknowledgements

We thank Vamsi and Ravindra Kini for their invaluable inputs and guidance.

References

1. People Who Are at Higher Risk for Severe Illness

2. covid_19 symptiom
3. Who is getting sick, and how sick? A breakdown of coronavirus risk by demographic factors

Add Comment

Post Comment

No Comments

BRIGHTIDEA