HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 3 (2024)

Dengue Virus Prediction Using Large Language Model

Srilalitha Madasu

Michael E. Debakey High School for Health Professions, USA

ABSTRACT

Dengue virus is a global issue most prominent in South American countries such as Brazil, Nicagua, Chile but can also be found in several other countries including Malaysia, Thailand, Taiwan, etc. Dengue virus has been around since the 1960's and is present to this day as there are nearly 400 million cases per year, which isn't including the suspected cases present in several rural areas. The problem is that many people do not have access to a doctor to give them medical aid. To help people without good healthcare facilities, the large language model comes in and serves as the doctor. The large language model predicts the virus that one has based on the symptoms that they are experiencing. Although there is no official vaccine for dengue virus, there are vaccines for other viruses such as Japanese Encephalitis, Malaria, Yellow Fever, etc. Based on the large language model's prediction of which virus one is experiencing, the appropriate treatment can be given to the patient.

Introduction

In several places throughout the world, Dengue Virus is prevalent and has been a problem since the 1960's. Some of these places include Brazil, Caribbean (specifically Puerto Rico), Malaysia, etc. There are about 100-400 million cases of dengue occurring every single year, however this may be way less compared to what is estimated as surveillance networks in rural areas aren't as good. Cases for Dengue fever have been going uphill every single year and will continue to go up unless a solution is come up to prevent any more cases. Dengue was a serious global issue up until there was an even bigger problem, the Covid-19 pandemic. Dengue fever diagnosis is a big challenge due to its similarities with other viruses such as Chikungunya, Zika, etc. Hence, a need for a diagnosis process and to make it easily accessible to the general public is vital. Hence, this paper developed a "Digital Doctor for Viruses" application to solve this issue.

Overview of Dengue Virus

How Dengue Virus Spreads

Dengue virus does not spread through human transmission, however it only spreads through a mosquito bite from the female Aedes Aegypti mosquito (1). Dengue Virus is an RNA virus from the Flaviviridae family of infections and an arthropod-borne virus (2). Dengue Virus primarily is more prevalent during the rainy season since mosquitoes love the moisture and humidity and like to lay their eggs during this time period, hence the mosquito population explosion while it rains (3). Many mosquitoes, including the Aedes Aegypti mosquito, like to lay their eggs in stagnant water like puddles, lakes, and ponds (3). Mosquitoes could also thrive from containers, which are not encapsulated, that contain water such as water bottlesEggs typically hatch in 7-10 days (3).

Symptoms of Dengue Virus

Dengue is also known as breakbone fever because of the seriousness of symptoms that occur because of this (4). Patients are in agonizing pain whilst having Dengue Fever (4). Symptoms typically appear 4-10 days after being infected and last for about 2-7 days (5). About one in four people experience mild symptoms of Dengue Virus as they have been infected for the first time (5).

Additionally, after 4 days of being infected with Dengue a person goes through Viremia, a condition l for having high amounts of the virus in the blood (6). Viremia lasts for about 5-12 days; approximately five days after being infected symptoms start to show, which last for a week or longer (6).

Symptoms include:

- Severe Headache (5)
- Joint Pains (5)
- Muscle Pains (5)
- Pain Behind The Eyes (5)
- High Fever (5)

However, for some people who have been re-infected with dengue, the symptoms will not be as mild as listed above (5). Severe infections such as Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS) most likely may occur (5). After 24-48 hours of the fever going away some more severe symptoms may occur (5).

Symptoms Include:

- Stomach Pain (5)
- Vomiting, which happens at least 3 times every 24hrs (5)
- Gum Bleeding (5)
- Nose Bleeds (5)
- Vomiting Blood (5)
- Blood is in the excretion (5)
- Tired, Restlessness and Irritability (5)

Risk

People of all ages living next to stagnant bodies of water, like lakes, ponds, and puddles are susceptible to getting Dengue Virus (5). However, children from ages 5-14 are more susceptible to Dengue virus followed by 15-24 year olds who have underdeveloped immune systems at the time (5). Older people, especially ones with pre-existing health conditions are even more prone to the virus (5).

Transmission and Vaccines

There are currently two Dengue Virus vaccines administered in Thailand which target all four of the dengue serotypes (5). Each of the Dengue vaccines is composed differently (5). Chimeric dengue vaccine in a yellow fever 17D backbone, is a vaccine given when an infant is 0 months old, 6 months old, and 12 months old (5). For people between 9-45 years old with a history of getting stricken by Dengue Virus must get the vaccine every 6 months (5). The second vaccine, Dengue tetravalent vaccine with dengue serotype 2 virus backbone, is given every three months to people between 4-60 years old, who have or have not been infected with Dengue Virus (5). Dengue Virus can be prevented only about 60%-80% and hospitalizations can be prevented 70%-90% using the vaccine (5). Vaccinations do not guarantee full protection from the virus (5). There is also no official vaccine for Dengue Virus (5).

Mortality

Approximately 400 million cases of infections occur per year and there is about 5%-20% mortality rate in areas (7). For example, in Malaysia dengue has become increasingly high (3). The number of cases in 1998 was 27,000 but that increased drastically in only 7 years when in 2005 they found out the cases were nearly 40,000 people (8). On a global scale there has been an increase of approximately 7 times when in 1990 the number of cases was 8.3 million but then in 2013 the cases rapidly increased to 58.4 million people (9). There was an especially high number of cases in the Western Hemisphere, particularly in tropical and subtropical regions such as Latin America, South-East Asia and the Western Pacific (9). Nearly 4 billion people live in areas where there is a risk of contracting Dengue Virus (7).

Using Data to Predict a Virus

Early and accurate detection of the dengue virus could save millions. Currently, we do not have an accurate and efficient method to diagnose the virus. Hence, a machine learning based model is developed to diagnose the virus. Specifically, a large language model is developed so that the patients can converse and use language to diagnose without the need to understand the details of the mathematical model. The conversational model uses Python programming language, Google large language model BARD (bard.google.com) , the graphical user interface is built in Streamlit. The model was trained on Google Colab for 6 hours. The data set for the model consists of 64 features such as sudden fever, headache, mouth bleed, muscle pain, joint pain, etc to predict 12 different viruses such as Chikungunya, Dengue Fever, Rift Valley Fever, Yellow Fever, Zika Virus, Malaria, Japanese Encephalitis, West Nile Fever, Plague, Tungiasis, and Lyme Disease. This data will be used to classify the virus and create a digital doctor using a large language model. The data to predict Dengue Fever has been taken from Kaggle. Fig 1 shows the statistical distribution of sudden fever and its correlation to other features and Fig 2 shows the feature headache correlation with other features. Fig 3 shows the data distribution for the different viruses. Fig 4 shows the correlation coefficient of the features wrt each other. Fig 5 shows the graphical user interface to select the symptoms to diagnose the virus. Fig 6 shows a sample of the code to converse with the large language model.

Below are the Statistical Plots of Sample Data of Dengue Virus Symptoms

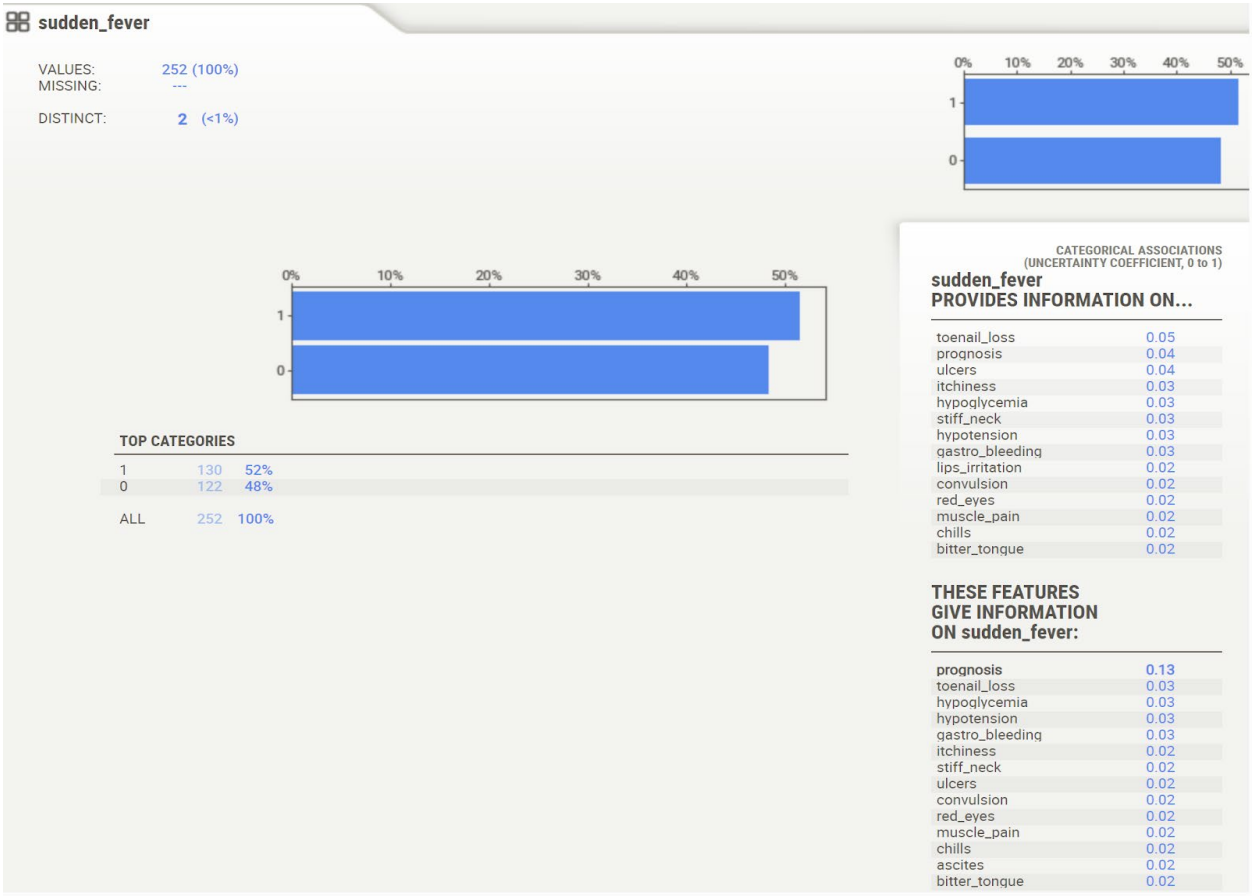

Figure 1. The image shows a bar plot for a symptom of Dengue Fever which is a sudden fever. The bar plot additionally shows how prevalent the symptom is, with people who were stricken by Dengue Virus. Data was derived from Kaggle.

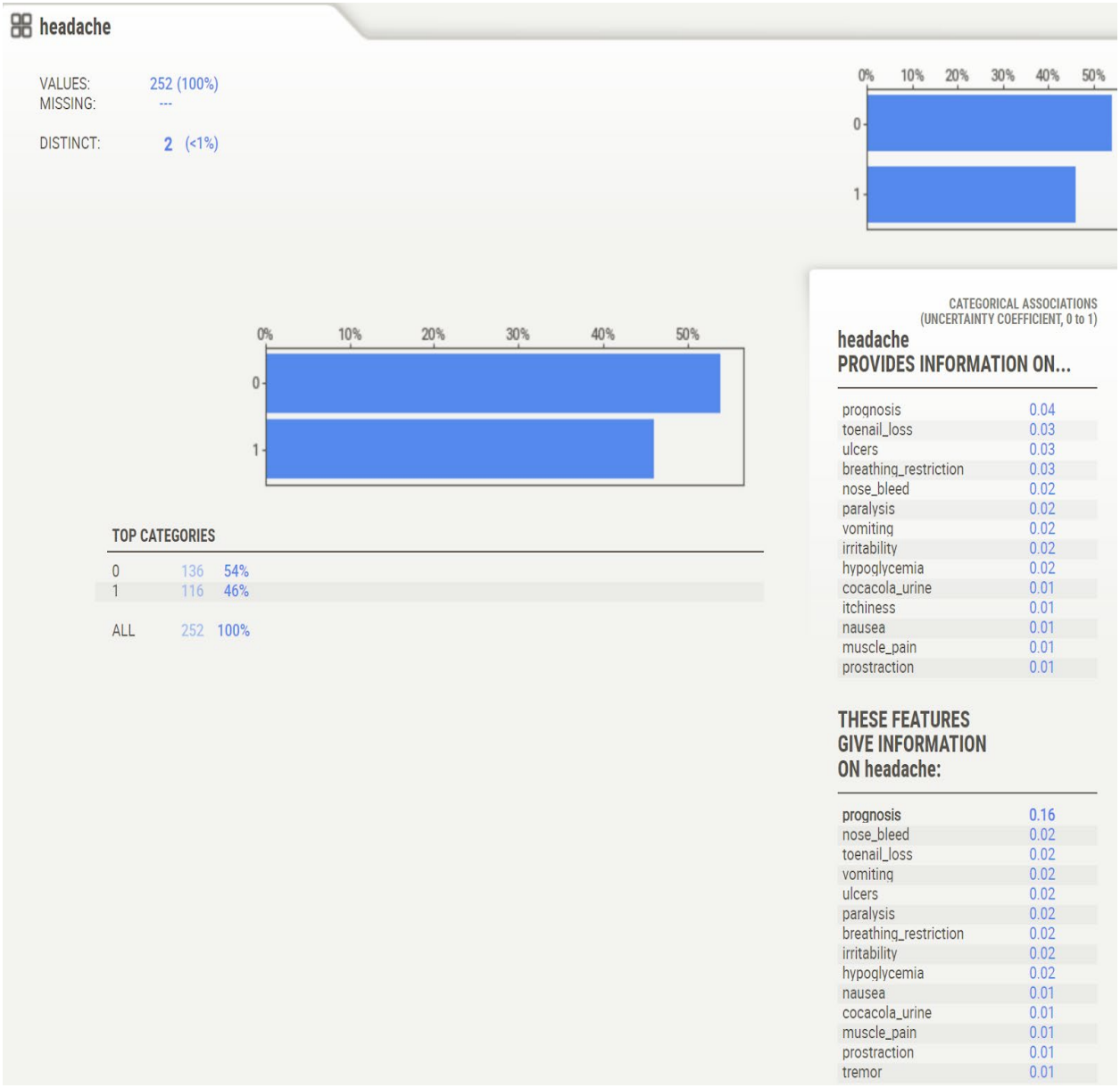

Figure 2. The image shows a bar plot for a symptom of Dengue Fever which is a headache. The bar plot additionally shows how prevalent the symptom is, with people who were stricken by Dengue Virus. Data was derived for Kaggle.

Figure 3. The bar plot shows how many types of viruses are going to be in the data set.

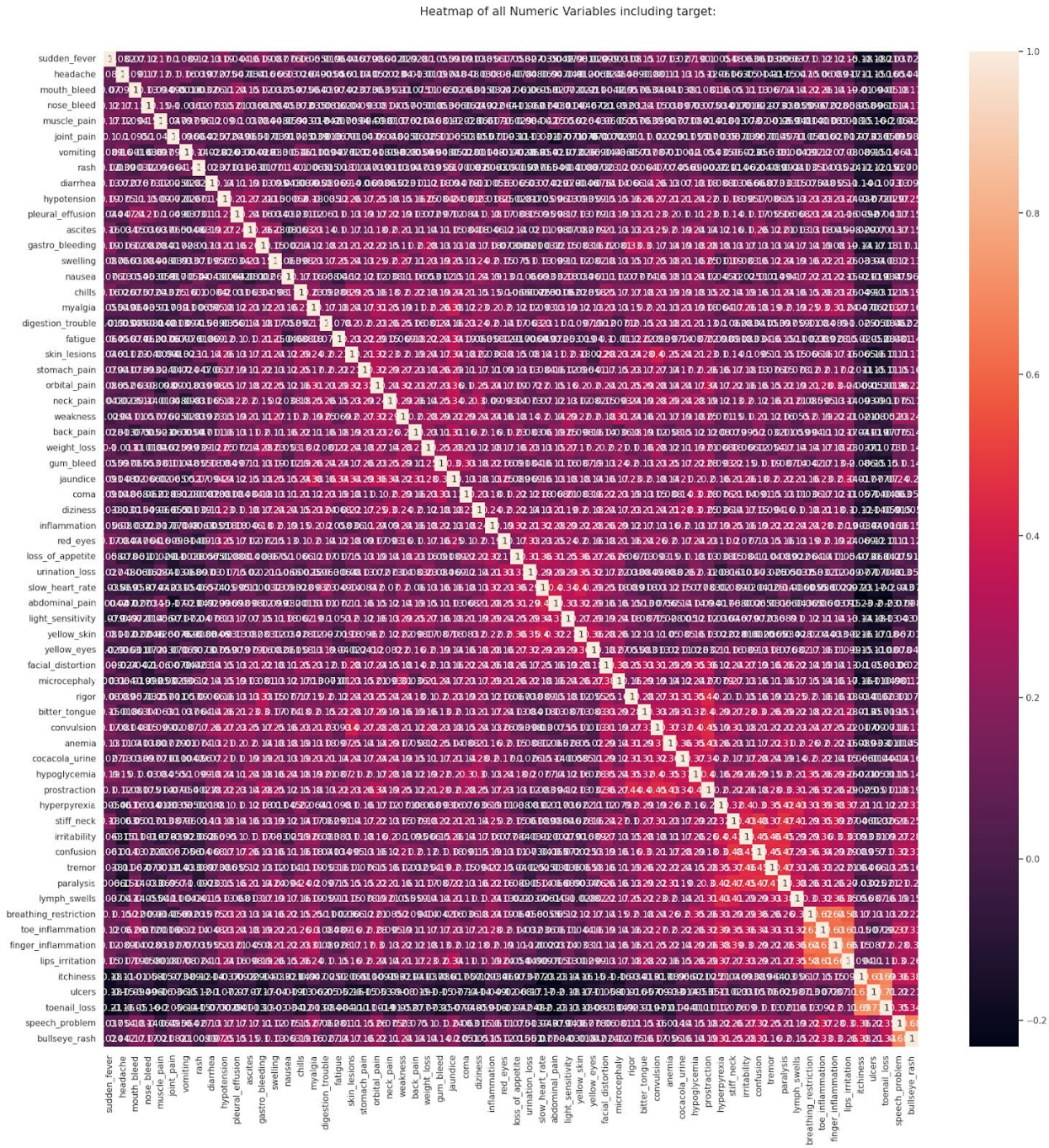

Figure 4. Correlation plot of the different symptoms of the different kinds of viruses. The closer it is to 1 then it is directly related, the closer it is to -1 then inversely correlated, the closer it is to 0 then no correlation with the symptom. Data was derived from Kaggle.

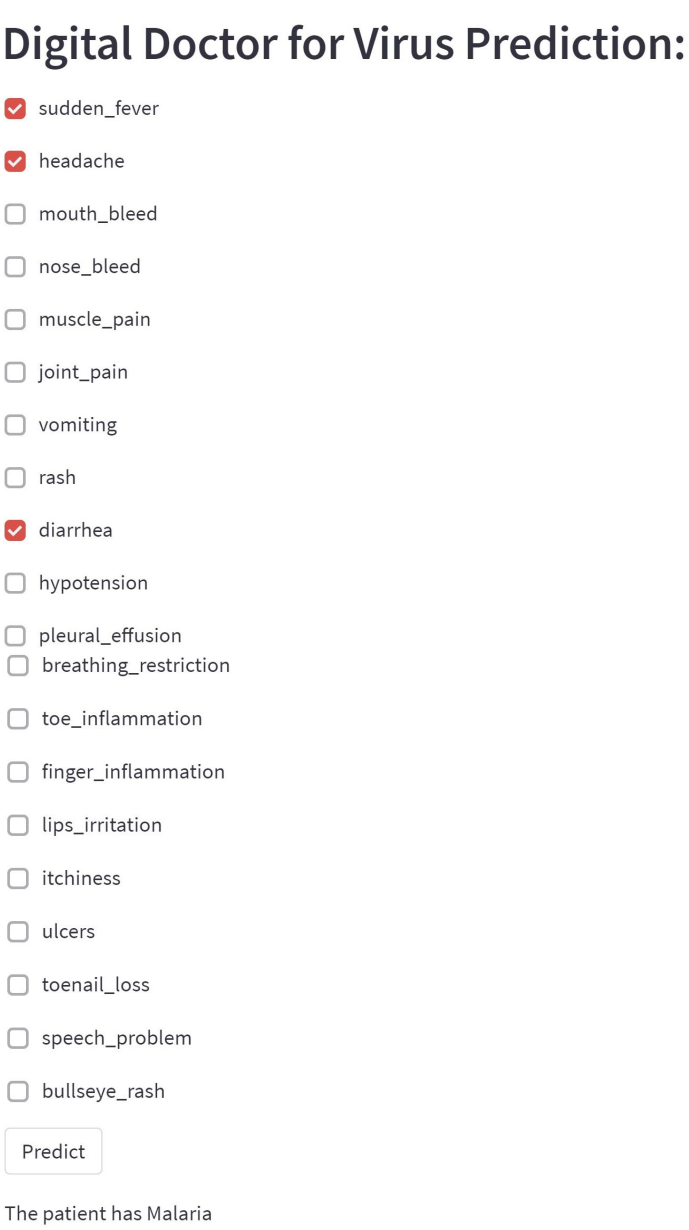

Figure 5. To test to see if the Large Language Model works, symptoms like sudden fever, headache, and diarrhea were selected. This experiment led to the diagnosis using Artificial Intelligence that the patient has Malaria. Hence proving that the digital doctor for viruses really works. Data driven from Kaggle.

```
# Choose a specific instance to explain
idx = 0  # Index of the sample in your dataset
text_instance = X_test.iloc[idx]['text']
pred = predictor(text_instance)
print(text_instance)
#print(pred)
#print(id2label[np.argmax(pred)])
print("The person has",id2label[np.argmax(pred)])
#print(text_instance)
```

```
This patient, doesn't sudden fever, has headache, doesn't mouth bleed, doesn't nose bleed, doesn't muscle pain, has joint pain, has vomiting,
The person has Plague
```

Figure 6. The image above shows the code written to come up with Fig 5. Data driven from Kaggle.

Conclusion

It is found that in areas with prevalent cases of dengue fever, there is not a proper treatment or medication being asserted to soothe the symptoms. This is due to the sole reason for poverty existing in such areas, so many cannot afford a doctor. A large language model has been developed for early and accurate diagnosis to get proper treatment for the virus one may be going through. 12 different viruses, such as Malaria, Dengue Fever, Japanese Encephalitis, Rift Valley Fever, Yellow Fever, etc, are able to be detected based on the symptoms like headache, nose bleed, vomiting, rash, diarrhea, joint pain, etc. Being able to detect a virus using Artificial Intelligence can save several millions of lives that are taken away from living in poor conditions such as Brazil which is seen to have the highest dengue rate out of the entire world. Saving lives because of Artificial Intelligence can not only reduce dengue to a great extent but also doesn't require much from the community to pay for, overall making the Large Language model a great alternative for quality healthcare.

Acknowledgment

I would like to thank Professor Katie for guiding me through the current vaccinations, mortality, risk, etc correlated with Dengue Virus. I would also like to thank Professor Virgel for helping me format my research paper. Additionally, I would also like to thank Dr. Srinath Madasu for helping me build the Large Language Model.

Limitations

Some challenges experienced while conducting research is building the large language model. While building the large language model, it was hard to train the model on a computer since the model has billions of parameters hence, Google Collab was used for the training of the model.

References

1. World Health Organization. (2019, October 24). *Dengue and severe dengue*. World Health Organization. https://www.who.int/news-room/questions-and-answers/item/dengue-and-severe-dengue#:~:text=Dengue%20cannot%20be%20read%20directly,reproduces%20in%20the%20blood%20system

2. Hasan, S., Jamdar, S. F., Alalowi, M., & Al Ageel Al Beaiji, S. M. (2016). *Dengue virus: A global human threat: Review of literature*. Journal of International Society of Preventive & Community Dentistry. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4784057/

3. *How to deal with insects after a storm.* How to deal with insects after a Storm. (n.d.). https://ortho.com/en-us/garden/how-to-deal-with-insects-after-a-storm.html#:~:text=Mosquitoes%20love%20moisture%20and%20 humidity,local%20mosquito%20population%20to%20explode.

4. Schaefer, T. J. (2022, November 14). *Dengue fever.* StatPearls [Internet]. https://www.ncbi.nlm.nih.gov/books/NBK430732/#:~:text=Dengue%20is%20a%20mosquito%2Dtransmitted,the%20usual%20duration%20of%20symptoms.

5. *Dengue vaccine, protecting yourself from the peak period for dengue fever.* Bumrungrad International Hospital. (2023, August 4). https://www.bumrungrad.com/en/health-blog/august-2023/dengue-vaccine-protecting-peak-period

6. Nature Publishing Group. (n.d.). Nature news. https://www.nature.com/scitable/topicpage/dengue-transmission-22399758/

7. Centers for Disease Control and Prevention. (2023, July 12). *Areas with risk of Dengue*. Centers for Disease Control and Prevention. https://www.cdc.gov/dengue/areaswithrisk/index.html

8. Lum, L. C., Sah, B. K., Shepard, D. S., Suaya, J. A., & Tan, L. H. (2008). Quality of life of dengue patients. *The American Journal of Tropical Medicine and Hygiene*, *78*(6), 862–867. https://doi.org/10.4269/ajtmh.2008.78.862

9. Tran, B. X., Thu Vu, G., Hoang Nguyen, L., Tuan Le Nguyen, A., Thanh Tran, T., Thanh Nguyen, B., Phuong Thi Thai, T., Latkin, C. A., Ho, C. S. H., & Ho, R. C. M. (2018, June 5). *Cost-of-illness and the health-related quality of life of patients in the dengue fever outbreak in Hanoi in 2017*. MDPI. https://www.mdpi.com/1660-4601/15/6/1174