|  |  |
| --- | --- |
| **Team Members** | |
| Fatemeh Khosravi (Ava) | 2743 9887 |
| Lachlan Muirden | 2875 9737 |
| Zaria Imran | 2777 1571 |

Project: Healthcare Support Chatbot

# Introduction

* Problem
* Importance
* Our motivation
* State input/output of model e.g. The input to our algorithm is an image, voltage of 10 sensors, etc.. We then use a CNN, LSTM, etc. to output a predicted age, decide which motor should be accelerated, etc.."

# Related works

* Existing papers
* Strengths and weaknesses
* Similarities/differences to our work
* What approaches were clever
* What is state-of-the-art

# Methodology

Skeleton Code Source: <https://github.com/python-engineer/pytorch-chatbot>

We have chosen to create a chatbot that is able to assist with first response healthcare, processing symptoms and giving diagnosis. The chatbot should also be aware of the likely severity of the collection of symptoms and provide recommendations based on this.

From our literature review, Kandpal et al. seemed the most basic implementation and feasible design. We have decided to use Natural Language Processing tokenizing, stemming and bag of words as described in all the literature we’ve explored. We are going to use our own model created with a Linear model with 4 hidden layers with dropout (but we would possibly increase if feasible) that fits our dataset while avoiding overfitting. If time permits, we will try to implement a model that remembers previous inputs (BRRN) such as that in Dhyani et al. Ideally we would have been able to use an XLNet and transfer learning, however this may take weeks and is probably unnecessary for our level of data.

# Dataset Description

* How many training/validation/test examples do we have
* Preprocessing required

Dataset Source: <https://www.kaggle.com/itachi9604/disease-symptom-description-dataset?select=symptom_Description.csv>

* + Symptom-severity.csv
  + dataset.csv
  + symptom\_Description.csv
  + symptom\_precaution.csv

This source contains 4 datasets, each of which will have utility for our chatbot. The first dataset, named ‘Symptom-severity’, contains a list of symptoms and each symptom’s associated severity, listed on a scale of one to ten. The second dataset, named ‘dataset’, contains a list of diseases and associated symptoms. The third dataset, named symptom\_Description, contains a list of illnesses, and associated descriptions of each illness. The fourth dataset, named symptom\_precaution, contains the same list of illnesses, but instead provides a list of precautions that can be taken if one is afflicted with each of the illnesses.

# Project Outcome Analysis

* hyper parameters
* how we chose them
* evaluation metrics
  + equations if necessary
* include quantitative and qualitative results
* visualisation of results
* examples of failure with our algoritm
  + discuss
* plots: legends, axis labels, and readable/clear font sizes

# Conclusions and Future Work

* summarise
* if more time, team members or computational resources, what would you explore?

# References

Graphical user interface, text, application

Description automatically generated