Final project

Name: Ngo Gia Han

StudentID: P76117022

# **Ideas and main flow**



Colleagues

Jobs

Best friends

family

Inspired from the social media community, which we usually have different groups of friends (such as family, colleagues, neighbors) but just display in general as our ***Friends.*** It’s hard to find an information related to a specific person that may belong to some group. For example, if we want to find his boss, it would be easier for us to seek the result in ***Jobs*** community, which will be more condensed and accurate

* Applied on COVID-19 subject

Query

Get one representative from each community

Detect community in the network

Build COVID-19 network

Using **Louvain** algorithm

Embed using **Distillbert COVID-QA** model

Weight of edge is relevance score between a pair of nodes

Using **Distillbert COVID-QA** model to build relationship between input paragraphs

Embedded query

Embedded documents

Get 3 clusters with highest similarity

Calculate similarity

# **Visualize network relationship**

Diagram

Description automatically generatedChart, diagram, bubble chart

Description automatically generatedA picture containing tree

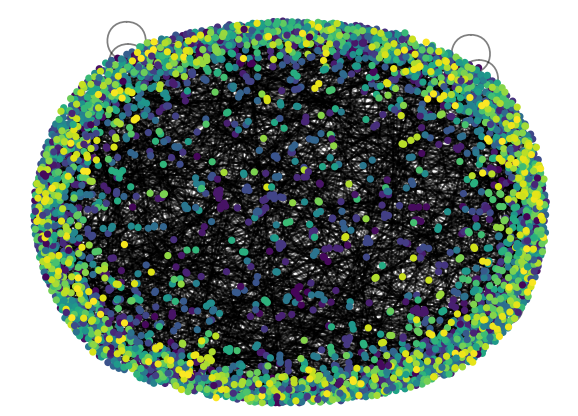
Description automatically generatedDataset: CORD-19 (3815 nodes, 2330 edges) 🡪 1519 communities

Figure 4: Example of a small community with relationship

Figure 3: Visualize by Gephi

Figure 1: Original network

Figure 2: Network after detecting community

For each node, keyphrase with 4 words is extracted to display the main content of that node