|  |  |
| --- | --- |
|  | **HASAN KALYONCU UNIVERSITY**  **Computer Engineering Department** **CENG 499 Project Proposal Form** |

**Part I. Project Proposer**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name Last name** | **Saed ALQARALEH** | **E-mail** | **saed.alqaraleh@hku.edu.tr** |
| **Company Information**  **(If you have collaboration with a company)** |  | | |

**Part II. Project Information**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Starting Term** | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 2 | 0 | 2 | 1 | / | 2 | 0 | 2 | 2 | |
| **Title of the Project** | CNN Based Classification System For Covid-19 Vaccine Related Tweets |
| **Project Description** | |
| Nowadays, a huge amount of information is shared through social media. This information could be opinions, feelings, complaints, requests, etc. Also, the Internet becomes a space where people feel free to share their thought with others. For example, currently due to the coronavirus, social media has an extremely large number of opinions, news about almost all the aspects of coronavirus. Such information is very useful either for governments or even societies. However, the extremely huge number of available information leads to the use of a very minimum amount of such useful information.  In this project, the performance of Deep Convolutional Neural Networks (ConvNet) will be first investigated and integrated to build an efficient and automated classification system for the Covid-19 vaccine-related tweets. The system can help to get the most relevant information about the people's opinions related to the vaccine. | |
| **Project Justification** | |
| **Novelty** | |
| **New aspects** | In our project, we will be creating a system that utilizes machine learning to efficiently analyze people's opinions about vaccines and make a decision whether the input sample(tweet) is positive, negative, or neutral. |
| **Complexity** | |
| **Challenging problem and issues** | The student will learn the principle of deep learning. |
| **Related computer science fields and subfields** | Artificial Intelligence (AI), Data Science, Machine Learning, Sentiment Analysis |
| **Tools** | Python and Python Libraries Pandas , Numpy etc. , Natural Language Toolkit (NLTK) |
| **Risk involved** | |
| **Potential problems and alternative solutions** | Incapability to deal with complex sentences and inadequate accuracy. As a solution, the performance of multiple state-of-the-art Machine Learning algorithms will be investigated in order to find the most suitable one. |
| **Minimum work required** | * Adequate knowledge and ability related to Natural Language Toolkit and Python programming language. * Group work |