**CENG 476 BIG DATA**

**PROJECT REPORT**

Mustafa Murat Hazneci 201711036

Eren Berk Saltaş 201711056

Ahmet Çetin Türkyener 201711066

Our aim in our project is to predict the **Diabetic Retinopathy** according to a specially taken picture of the person's eye. We fed the existing data to the machine learning algorithm and trained a model, from the model expected to predict according to the new data.

We preferred mongodb database to retrieve data for processing in our project because mongodb is fast, It is open source, able to hold unstructured data. To extract and process data from mongodb we needed to connect with our own environment, we used the very compatible pymongo library for this Then, with our machine learning application, we predicted the disease according to the image in the data from mongo. Our project allows us to make batch and real time disease prediction and we wrote the results to PostgreSQL. We kept the data according to the disease class in the tables in PostgreSQL and expanded our dataset so that. In the future, by using the new data, our machine learning algorithm will learn more.