

Invitation to M.Tech. Thesis Defense of Divya Sharma: February 12, 2020 (Wednesday): 15.00-16.00 IST

In Partial Fulfillment of the Requirements for the Degree of

M.Tech. CB

Divya Sharma (MT18235)

Will defend her thesis

Title: "Automated Discovery of Abstracts Reporting Gene-Disease Relationship"

IIIT-D Faculty and Students are invited

Date: February 12, 2020 (Wednesday)

Time: 15.00-16.00 IST

Place: CB Meeting Room (3rd Floor, R&D Building)

Examiner:	Internal:	Gaurav Ahuja
	External/Internal:	Arjun Ray
	Advisor:	Debarka Sengupta

Abstract

Text classification is a construction problem of models which can classify new documents into predefined classes. It is important before text mining that we know what is the most important data that we require for our research. Text mining has become an essential tool for biomedical research. Our project aims to identify the gene-disease relationship using natural language processing techniques and word embeddings. Assignment of high-dimensional vectors (embeddings) to words in a text corpus in a way that preserves their syntactic and semantic relationships is one of the most fundamental techniques in natural language processing (NLP). We present a completely generic model based on statistical word embeddings, which shows the gene similarity and proves the gene-disease relationship using word analogies.

