

**Invitation to M.Tech. Thesis Defense of Shivanshu Kukrety: June 13, 2022 (Monday): 11:00 AM-11:45 AM IST**

In Partial Fulfillment of the Requirements for the Degree of

**M.Tech. CB**

**Shivanshu Kukrety (MT20217)**

Will defend his thesis

**Title: “Blood based prediction of immune mediated disorders from bulk RNA-seq data”**

IIIT-D Faculty and Students are invited

**Date: June 13, 2022 (Monday)
Time:** **11.00 AM-11.45 AM IST**

**Meeting Link(G-Meet):** [**https://meet.google.com/zts-uzdi-xyx**](https://meet.google.com/zts-uzdi-xyx%20)

**Examiner: Internal:   Gaurav Ahuja**

**~~External~~/Interna~~l~~: N Arul Murugan**

**Advisor: Debarka Sengupta**

**Co-Advisor Jeremie Poschmann (Nantes University, France)**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Abstract**

Immune-mediated disorders (IMDs) include a wide spectrum of pathologies ranging from autoimmunity to autoinflammation, and they impact a substantial number of individuals worldwide. Although dysfunctional inflammatory cytokine behaviour in IMDs implies abnormal immune cellular activity, not much is known about the underlying responsible genes and sometimes even crucial cell types. Moreover the proportions of gene expression variance explained by the clinical diagnosis is quite small, which makes it difficult to analyze the underlying condition. Recent breakthroughs in artificial intelligence have resulted in widespread industrial and academic use, with machine learning systems outperforming traditional schemes in a wide array of applications. Our project aims to make use of this predictive power to build classification models using gene expression data for prediction of immune mediated diseases. The various models built are tested using nested cross validation on a wide variety of metrics to analyze the generalizability of our classifiers. The best result was achieved by support vector machines with an accuracy of 92.29% and a MCC value of 91.59%. In our project we have also carried out a differential expression analysis in order to obtain a comparison of gene expression patterns between a healthy individual and a patient infected with an immune mediated disease, which enable us to identify genes which may be participating in specific functions such as protein synthesis, hormone delivery and pathological pathways. In addition to the ones stated we also present a deconvolution operation performed on cibersort to obtain the relative cell proportions of 28 immune cell types in a specific disease class. The codebase is available on github and can be accessed using the following link: https://github.com/tom8861/thesis.