

**Invitation to M.Tech. Thesis Defense of Dinesh Joshi: June 13, 2022 (Monday): 09:30 AM-10:15 AM IST**

In Partial Fulfillment of the Requirements for the Degree of

**M.Tech. CB**

**Dinesh Joshi (MT20208)**

Will defend his thesis

**Title: “Transfer learning and consensus clustering-based analysis of single-cell data”**

IIIT-D Faculty and Students are invited

**Date: June 13, 2022 (Monday)
Time:** **09.30 AM-10.15 AM IST**

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

**Examiner: Internal:   Jaspreet Kaur Dhanjal**

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

**Advisor: Debarka Sengupta**

**Co-Advisor NA**

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

**Abstract**

With the rapid development of sequencing technologies, single-cell genomics has become a primary focus of researchers. Through single-cell RNA-sequencing data analysis, it has now become possible to study some fundamental questions in biology at the celllevel resolution like which cell types are present in a tissue, what function these cells carry and how are these functions different from healthy tissues etc. However, utilizing publicly available large scale cell atlases for analysis is not so easy and has its own challenges. Since single-cell data analysis workflow comprises many steps like data harmonizing, batch effect correction, normalization, visualization, clustering, cell-type classification and differential gene expression analysis, so different methods have to be used to carry out these tasks. But with advancements in computational power and techniques, utilising large scale reference datasets to gain knowledge and then transfer it to smaller query datasets has become common. So to facilitate the users with these recent transfer learning (TL) techniques to do single-cell analysis, we have come up with Transcend, a webserver which hosts a number of such pretrained TL methods. We hope that our webserver will be widely adopted as a single source for exploring transfer learning methods for single-cell analysis.

Moreover, clustering single-cell RNA-seq data to uncover patterns, can lead to the identification of new or rare cell types or subtypes but clustering single-cell data is also a challenge due to the highly sparse read count data and due to presence of technical variation in the data it becomes difficult to capture biological variation in data, which can lead to poor clustering of cells. So to get better clustering results, we explored the concept of consensus clustering. After trying out multiple consensus clustering algorithms, we were able to get improved clustering results which are stable and can be further used to improvise the clustering results of other methods.