

August 29, 2025

### Acknowledgment Letter for Mr. Swaroop M. Ratnaparkhi

This document serves to acknowledge the ongoing Machine Learning project undertaken by Mr. Swaroop Mahadeo Ratnaparkhi (Roll No.: 20CY20035) on 19th July 2025, under the guidance of **Dr. Ľuboš Buzna**, Faculty of Management Science and Informatics, Department of Mathematical Methods and Operations Research, Associate Professor of Applied Informatics.

Mr. Ratnaparkhi is currently in the final year of the 5-Year Integrated M.Sc program in the Department of Chemistry at the Indian Institute of Technology Kharagpur.

- **Project Overview**

This project focuses on developing machine learning models trained on fluorescence spectra obtained from urine samples of patients diagnosed with various forms of cancer. The primary objective is to enable predictive classification—determining whether a patient is likely to be suffering from cancer based on their urine fluorescence data.

- **Work Completed to Date**

1. Data Preprocessing:

- Performed data cleaning and applied Savitzky-Golay smoothing.

2. Binary Classification: Control vs. Cancer Group

- Logistic Regression with L1 Regularization: Implemented to perform initial classification and identify key predictive features. These selected features were subsequently used to train a standard logistic regression model (without regularization) for further classification.

- Lasso Regression for Feature Selection: Used to extract significant features from the dataset. Classification thresholds were defined to distinguish between non-cancerous (-1) and cancerous (+1) groups. A logistic regression model (without penalty) was then trained on the significant features.

3. Model Optimization

- Conducted hyperparameter tuning using GridSearchCV.

- Applied oversampling techniques to address class imbalance and improve predictions.

- Adjusted classification thresholds to optimize the conversion of predicted probabilities into discrete class labels.

4. Multiclass Classification

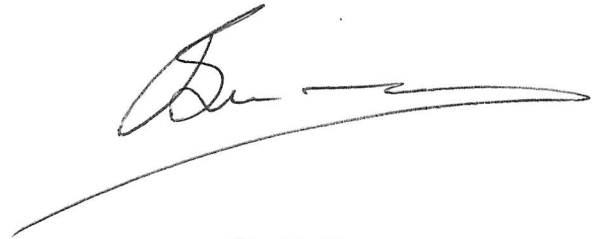
- Implemented One-vs-Rest approach using Linear Regression for basic multiclass classification.
- Extended the multiclass framework using Logistic Regression (One-vs-Rest).
- Implemented One-vs-One approach using Logistic Regression for Multiclass classification.
- Trained Decision Trees and Random Forest for Multiclass classification.

*A major challenge in this work is the limited availability of data and significant class imbalance.*

**The project is currently ongoing, and Mr. Ratnaparkhi has decided to continue to work.**

Kindly,

*In Žilina 29/03/ 2025*



Dr. Ľuboš Buzna,

Associate Professor of Applied Informatics,

Department of Mathematical Methods and Operations Research,

Faculty of Management Science and Informatics,

University of Žilina

Žilina Region

Slovakia

ŽILINSKÁ UNIVERZITA V ŽILINE  
FAKULTA RIADENIA A INFORMATIKY  
Univerzitná 8215/1  
010 26 Žilina  
-1-