# VISVESVARAYA TECHNOLOGICAL UNIVERSITY BELAGAVI -590018



**A TECHNICAL SEMINAR REPORT ON**

**MACHINE LEARNING IN BIOINFORMATICS: NEW TECHNIQUE FOR DNA SEQUENCING CLASSIFICATION**

**A Technical Seminar Report Submitted in Partial Fulfillment of the Requirements for the VIII Semester B.E**

**Submitted By**

**Rajashekhar Naduvinahalli**

**2KA21CS037**

### Under the Guidance of

**Dr. Arunkumar Joshi**

**Associate Professor**

****

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING SMT. KAMALA & SRI VENKAPPA M. AGADI**

**COLLEGE OF ENGINEERING& TECHNOLOGY LAXMESHWAR 2024-25**

****

**Smt. Kamala & Sri Venkappa M. Agadi**

**College of Engineering and Technology,**

**Laxmeshwar-582 116**

***Certificate***

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

This is to **certify that** Mr. Rajashekhar Naduvinahalli bearing the **USN 2KA21CS037**,have satisfactorily completed the Technical Seminar entitled “**Machine Learning in Bioinformatics: New Technique for DNA Sequencing Classification”** in partial fulfillment for the award of the degree of Bachelor of Engineering of Visvesvaraya Technological University Belagavi, during the year 2024-25. Technical Seminar Report has been approved, as it satisfies the academic requirements in respect of Technical Seminar Work prescribed for the said degree.

………………. ……………….…….

**Seminar Guide Examiner**

**Dr. Arunkumar Joshi Prof. Prakash Hongal**

……………….……. …………………. ………………….

**Seminar Coordinator HOD Principal**

**Dr. Arunkumar Joshi Dr. Arun Kumbi Dr. Parashuram Baraki**

**Acknowledgment**

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encouragement crowned out efforts with success.

I would like to take this opportunity to thank my Technical Seminar Guide **Dr. Arunkumar Joshi**, Associate Professor Department of Computer science and Engineering, without his immense guidance and support the work would have been unthinkable, helped me in the completion of technical seminar work.

I express my deep sense of gratitude to our HOD **Dr. Arun Kumbi**, Department of Computer science and Engineering, for his unstinted support.

I extend my gratitude to the Principal Dr. Parashuram Baraki, SKSVMACET, Laxmeshwar for the generous support in all regards. I extend my heartfelt thanks to all the faculty members, teaching and nonteaching staff of department of Computer science and Engineering, SKSVMACET, Laxmeshwar who have helped me directly or indirectly. I’m very much indebted to my parents and friends for their unquestioning best cooperation and help

### Mr. Rajashekhar Naduvinahalli

**2KA21CS037**

### ****List of Abbreviations****

DNA Deoxyribonucleic Acid

PCR Polymerase Chain Reaction

ML Machine Learning

NLP Natural Language Processing

ANN Artificial Neural Network

FCGR Frequency Chaos Game Representation

SVM Support Vector Machine

TP True Positive

TN True Negative

FP False Positive

FN False Negative

CAD Coronary Artery Disease

SNP Single Nucleotide Polymorphism

TPOT Tree-based Pipeline Optimization Tool

MFN2 Mitofusin 2 (Mitochondrial Fusion Protein 2)

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats

### ****Abstract****

In recent years, the integration of machine learning into healthcare has revolutionized disease detection, offering significant advancements in diagnostic accuracy, efficiency, and accessibility. This project explores the application of machine learning techniques in disease detection by analyzing diverse medical datasets, including patient records, clinical imaging, and symptom-based data. The study investigates various machine learning models, including supervised, unsupervised, and deep learning approaches, to identify the most effective methods for early disease diagnosis.

The methodology involves data preprocessing, feature selection, model training, and evaluation using performance metrics such as accuracy, precision, recall, and ROC-AUC scores. The study also emphasizes the importance of ethical considerations, including data privacy, security, and bias mitigation, to ensure reliable and fair decision-making in medical applications.

The findings highlight the potential of machine learning in automating disease detection, thereby reducing human error and enabling timely intervention. The project aims to contribute to the field of precision medicine, enhancing patient care and supporting medical professionals with decision-making tools. Future research will focus on improving model interpretability, integrating real-time monitoring systems, and expanding datasets for better generalization.

|  |  |  |
| --- | --- | --- |
| Chapter | Title | Page No |
| 1 | Introduction | 1-4 |
| 2 | Literature Survey | 5-6 |
| 3 | Methodology | 7-12 |
| 4 | Result and Discussion | 13-15 |
| 5 | Applications and Challenges | 16-19 |
|  | Conclusion and future scope | 20 |
|  | Reference | 21-22 |

**CONTENT**