

# Developing a Machine Learning library in C++

Tanay Shinde<sup>1</sup>, Kaushik Nayanala<sup>2</sup>, Ujjawal Agarwal<sup>3</sup>, Twisha Goyal<sup>4</sup>

<sup>1,2,3,4</sup>B.Tech, MIT ADT

## Abstract:

This project aims to develop a comprehensive and versatile machine learning library in C++ tailored to address the diverse needs of developers and researchers in the field. The library encompasses a robust set of core machine learning algorithms, encompassing supervised, unsupervised, and reinforcement learning techniques. Additionally, it incorporates essential data preprocessing tools to streamline data manipulation and feature engineering tasks, along with model evaluation capabilities crucial for assessing algorithm performance. The library's primary focus is on providing a rich suite of machine learning algorithms. This empowers users to effectively prepare data for training machine learning models. Additionally, the library provides tools for data splitting, and model evaluation to ensure reliable and robust model performance assessment.

**KEYWORDS:** Supervised, Unsupervised, Model evaluation, Data Splitting, Performance Assessment

## 1. INTRODUCTION (*HEADING 1*)

The study of social influence within networks has garnered significant attention, particularly concerning the behavior of communities and individuals in influencing information spread. Several researchers, such as Kempe et al. (2003), Bharathi et al. (2007), Bonchi (2011), Domingos & Richardson (2001), and Leskovec et al. (2007), have investigated influence maximization and social diffusion models, aiming to identify nodes in networks that can effectively enhance information propagation. To develop a machine learning library in C++ for this purpose, the focus would be on integrating graph theory, network analysis, and various machine learning algorithms. The library would encompass functions for community detection, identifying influential nodes through measures like PageRank, modeling homophily effects, and potentially employing machine learning models to predict or classify social influence behaviors. By combining these methodologies, the library aims to provide tools for the comprehensive analysis of social influence, aiding in understanding the dynamics of information spread within networks and its implications for marketing strategies or information diffusion. This restructured summary provides an overview without detailing specific points but encapsulates the essence of building a machine learning library in C++ for studying social influence in networks.

## 2. PROBLEM STATEMENT

Machine learning has become an integral part of various industries, but the availability of high-performance machine learning libraries in C++ is limited. This project aims to address this issue by developing a comprehensive machine learning library in C++. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads- the template will do that for you.

### 3. OBJECTIVE

This project focuses on designing and implementing a versatile machine learning library in C++ to cater to the needs of developers and researchers in the field. The library will include core machine learning algorithms, data preprocessing tools, and model evaluation capabilities.

**Table 1. Survey of project**

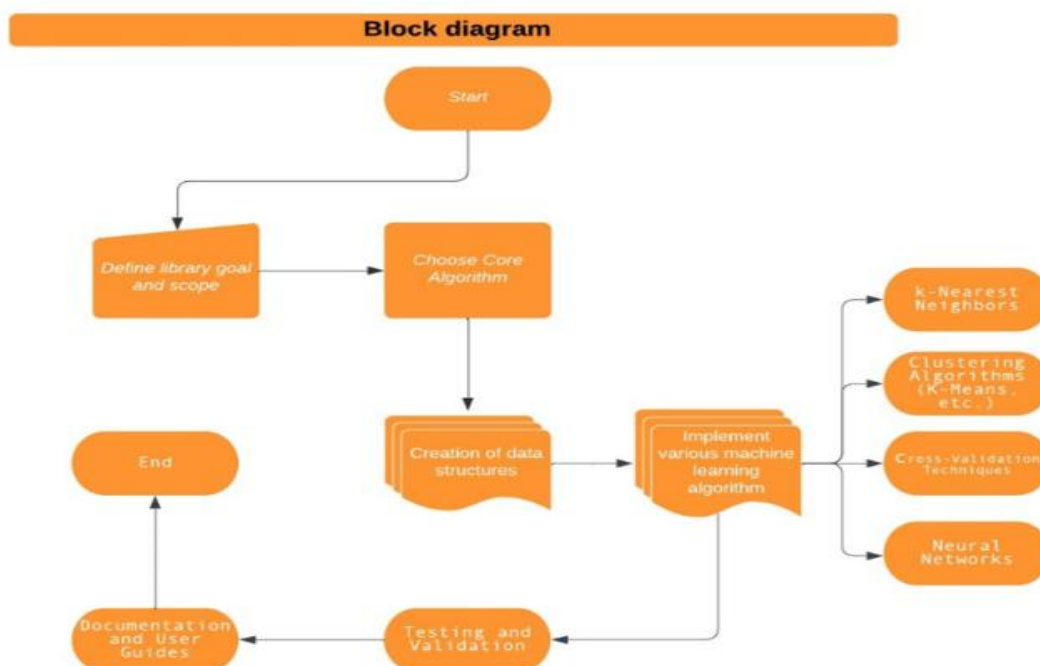
S.r.No.	Title of the page	Author name	Remark
1	ML PACK:A Scalable C++ ML Library	Ryan R. Curtin	Goal of ML library like this ML PACK is to combine efficiency and accessibility
2	Hands on ML with C++	Krill koladiashnyi	Pre-processing in C++

### 4. PROPOSED SYSTEM

Our proposed system will be a comprehensive machine learning library in C++, designed to address the limitations of existing options. Key features include:

1. Core Algorithm: Covering supervised, unsupervised, and reinforcement learning.
2. Efficient Preprocessing: Streamlined data cleaning, normalization, and feature selection.
3. Model Evaluation: Providing metrics for assessing model performance.
4. Scalability: Leveraging C++ for high-performance implementations.
5. Integration: Designed for seamless integration with existing C++ frameworks.

This system aims to empower developers with efficient, scalable, and easily integrated machine learning capabilities in C++.



## 5. ACKNOWLEDGMENT

WE SINCERELY WISH TO THANK OUR PROJECT GUIDE PROF. KOMAL MUNDE FOR HER ENCOURAGING AND INSPIRING GUIDANCE THAT HELPED US TO MAKE OUR PROJECT A SUCCESS. OUR PROJECT GUIDE MADE SURE WE WERE ON TRACK AT ALL TIMES WITH HIS EXPERT GUIDANCE, KIND ADVICE, AND TIMELY MOTIVATION WHICH HELPED US DETERMINE OUR PROJECT. WE ALSO EXPRESS OUR DEEPEST THANKS TO OUR HOD DR. SHRADDHA PHANSALKAR WHOSE BENEVOLENCE HELPED US BY MAKING THE COMPUTER FACILITIES AVAILABLE TO US FOR OUR PROJECT IN OUR LABORATORY AND MAKING IT A TRUE SUCCESS. WITHOUT HER KIND AND KEEN CO-OPERATION, OUR PROJECT WOULD HAVE BEEN STIFLED TO A STANDSTILL. LASTLY, WE WOULD LIKE TO THANK OUR COLLEGE PRINCIPAL FOR PROVIDING LAB FACILITIES AND PERMITTING US TO GO ON WITH OUR PROJECT. WE WOULD ALSO LIKE TO THANK OUR COLLEAGUES WHO HELPED US DIRECTLY OR INDIRECTLY DURING OUR PROJECT.

## 6. CONCLUSION

In conclusion, our project to develop a Machine Learning Library in C++ fills a crucial gap in the field. Our system offers core algorithms, efficient preprocessing, model evaluation, scalability, and seamless integration.

By leveraging C++'s performance, we empower developers to innovate across industries. We invite collaboration to ensure our system meets evolving needs, driving groundbreaking applications and discoveries.

Thank you for joining us on this journey.

## REFERENCES

1. John D. Kelleher, Brian Mac Namee, and Aoife D'Arcy, "Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies," The MIT Press, 2015.
2. Trevor Hastie, Robert Tibshirani, and Jerome Friedman, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction," Springer, 2nd Edition, 2009.