WASML: Dataframes, Preprocessing and Clustering

Pinto, Dion Bhonsle, Archit Bhonsle, Archit

April 21, 2021

Contents

Abstract

Machine learning is an integral part of the modern business as well as tech consumers. It is utilized by various organizations to improve their products for better object detection, speech detection etc. Many organizations now run ML models on the browser e.g. recommendation of videos.The most accessible way to interact with machine learning is through the browser. Tensorflow Js is a perfect example of using machine learning on the web, However, TFJS utilizes JavaScript for its implementation which is much slower when compared to a compiled language like C++/Rust. This project aims to create a fast machine learning library on the web which implements features of ML libraries and runs them at near native speed on the browser. It will build on modules such as a Data Analysis and Manipulation Tool, in which we can ingest data and manipulate it, process the datasets, Apply the ingested data on ML algorithms for clustering , reducing the dimensionality of the dataset etc all utilizing Rust for greater performance on the browser.