

ML Tool - Multi-purpose Machine Learning Tool



Full demo of the application:

<https://www.dropbox.com/s/h11afe4xx6xwqtu/ML%20Tool%20Demo%20Video.mp4?dl=0>

A web application tool for

- Data Pre-Processing
- Data Visualization
- ML Model Training

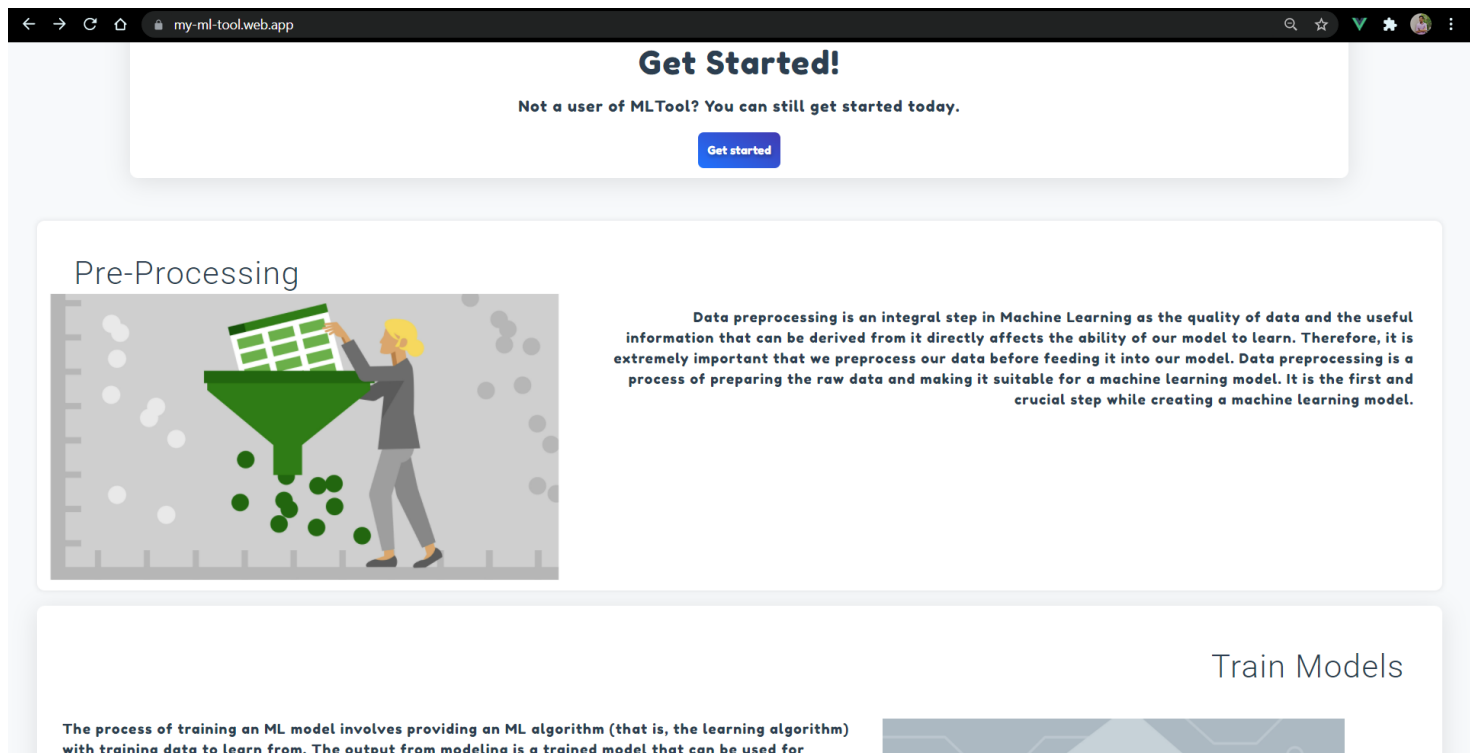
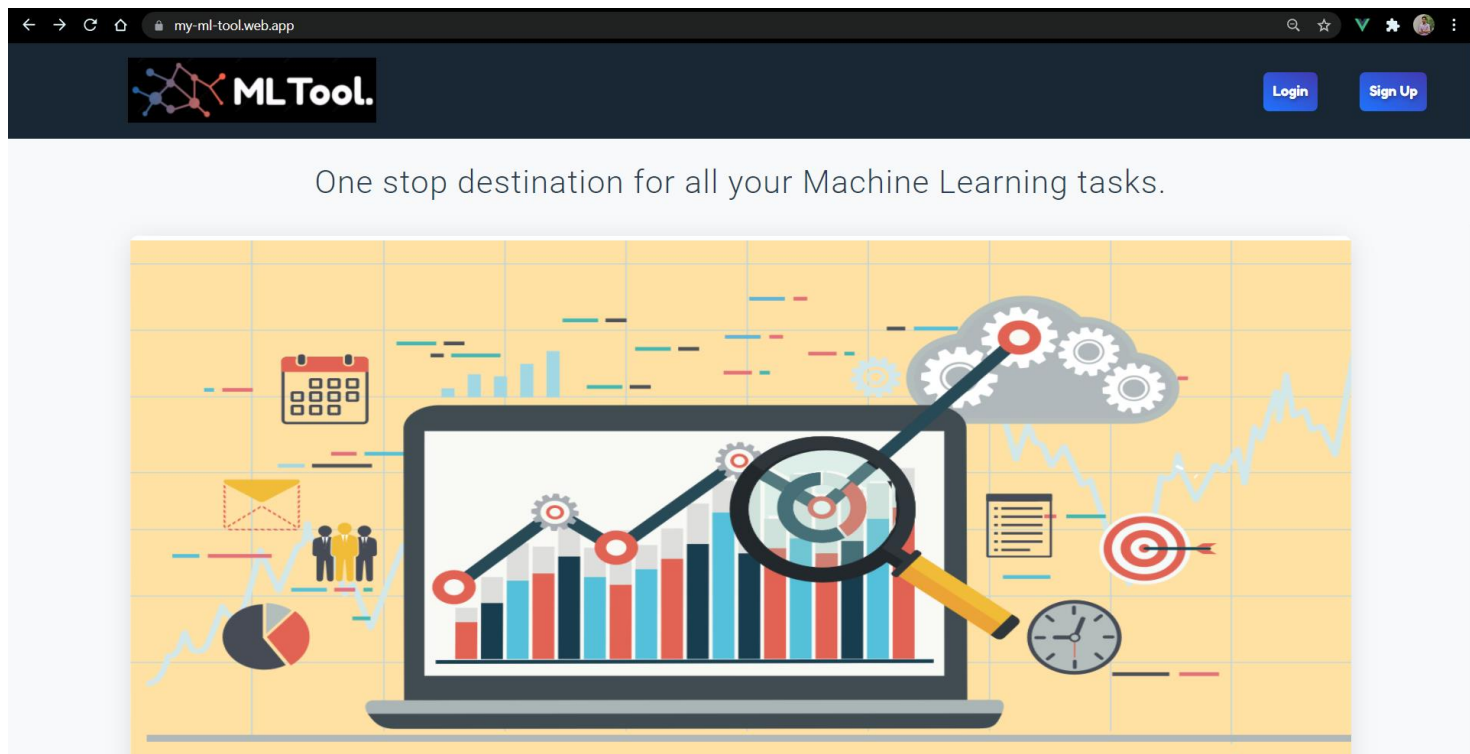
Frameworks used:

- Front-end Framework: Vue JS
- Back-end Framework: Flask

Novel Features:

- Option for users to pre-process the dataset automatically (if user is not sure how to pre-process data) or customise the attributes for pre-processing.
- Get best possible parameters for the ML Model. It uses Randomized Search Algorithm to get parameters using which best possible accuracy for a model can be achieved.

Homepage:



Train Models

The process of training an ML model involves providing an ML algorithm (that is, the learning algorithm) with training data to learn from. The output from modeling is a trained model that can be used for inference, making predictions on new data points.



Visualization



Data visualization is a technique that uses an array of static and interactive visuals within a specific context to help people understand and make sense of large amounts of data. The data is often displayed in a story format that visualizes patterns, trends and correlations that may otherwise go unnoticed. Data visualization is regularly used as an avenue to monetize data as a product.

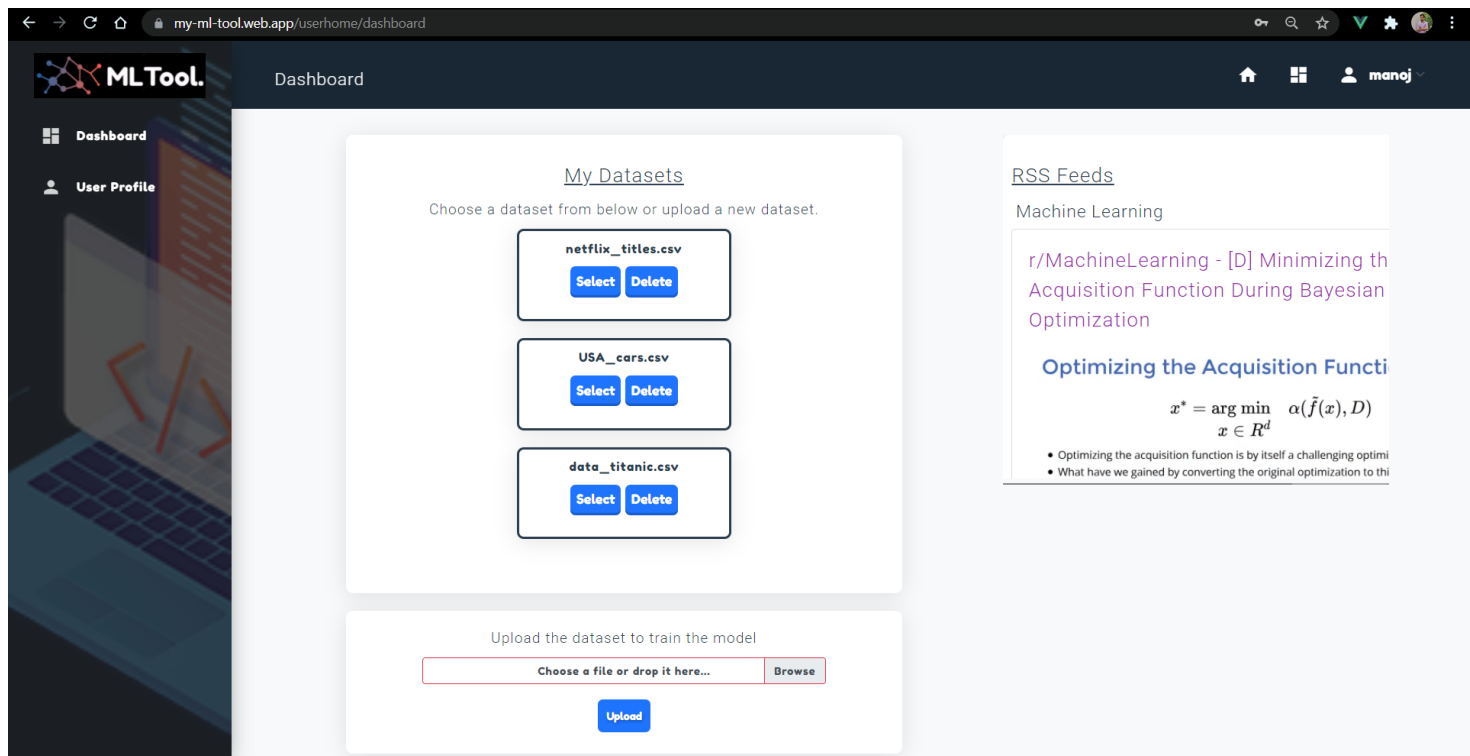
ML Information

Analytical reporting has been evolving into one of the most important business intelligence components that has compelled companies to adapt their strategies based on data-driven insights. While many companies struggle to leverage an effective business intelligence strategy, the importance of analytical information created fluctuation of data that cannot be simply collected into a single spreadsheet. It has become harder to create and use a single report and communicate a wide range of vital information between departments, stakeholders and important parties in a single company.



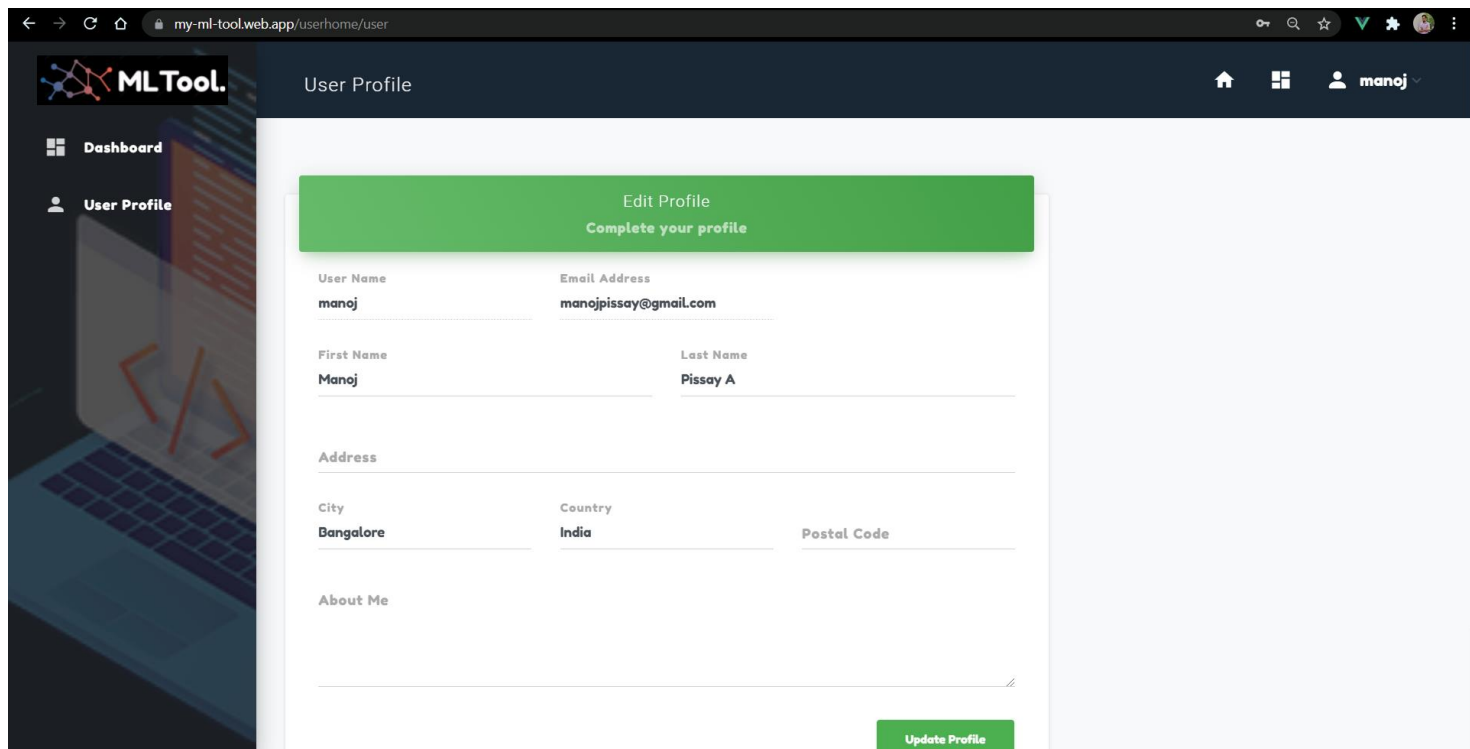
Scroll to top ^

Dashboard:



The screenshot shows the MLTool Dashboard. The browser address bar displays `my-ml-tool.web.app/userhome/dashboard`. The left sidebar contains the MLTool logo and navigation links for 'Dashboard' and 'User Profile'. The main content area is titled 'Dashboard' and features a 'My Datasets' section with the instruction 'Choose a dataset from below or upload a new dataset.' Below this, three dataset cards are listed: 'netflix_titles.csv', 'USA_cars.csv', and 'data_titanic.csv', each with 'Select' and 'Delete' buttons. At the bottom of the dashboard, there is an upload section with the text 'Upload the dataset to train the model' and a file input field with a 'Browse' button and an 'Upload' button. On the right side, there is an 'RSS Feeds' section titled 'Machine Learning' containing an article snippet about 'Minimizing the Acquisition Function During Bayesian Optimization' with the title 'Optimizing the Acquisition Function' and a mathematical equation
$$x^* = \arg \min_{x \in R^d} \alpha(\tilde{f}(x), D)$$
.


User Profile:



The screenshot shows the MLTool User Profile page. The browser address bar displays `my-ml-tool.web.app/userhome/user`. The left sidebar is identical to the dashboard view. The main content area is titled 'User Profile' and features an 'Edit Profile' section with the heading 'Complete your profile'. The form includes fields for 'User Name' (manoj), 'Email Address' (manojpissay@gmail.com), 'First Name' (Manoj), and 'Last Name' (Pissay A). There is also a single-line 'Address' field. Below these, there are fields for 'City' (Bangalore), 'Country' (India), and 'Postal Code'. An 'About Me' section with a text area is located at the bottom of the form. A green 'Update Profile' button is positioned at the bottom right of the form.

Dataset Pre-processing:

my-ml-tool.web.app/features/preprocessing

Preprocessing

manoj

Data Pre-Processing for netflix_titles.csv

Target Column:

Please select

Not sure how to pre-process the data?

Pre-Process Data

Features to remove:

☐ show_id

☐ type

☐ title

☐ director

☐ cast

☐ country

☐ date_added

☐ release_year

my-ml-tool.web.app/features/preprocessing

manoj

☐ listed_in

☐ description

Convert Catagorical data to numeric data using:

Please select

Replace the below Columns having NaN with:

cast:

Please select

country:

Please select

date_added:

Please select

director:

Please select

rating:

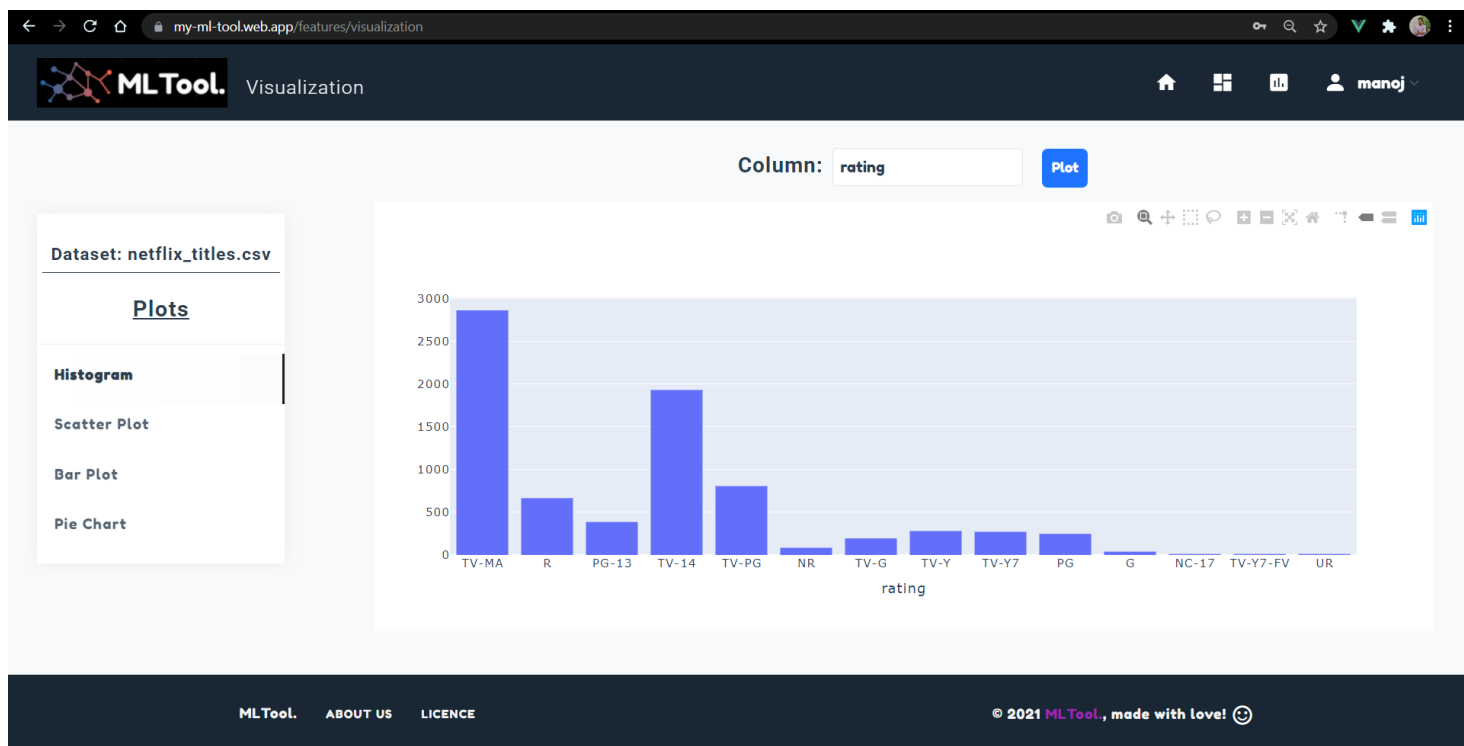
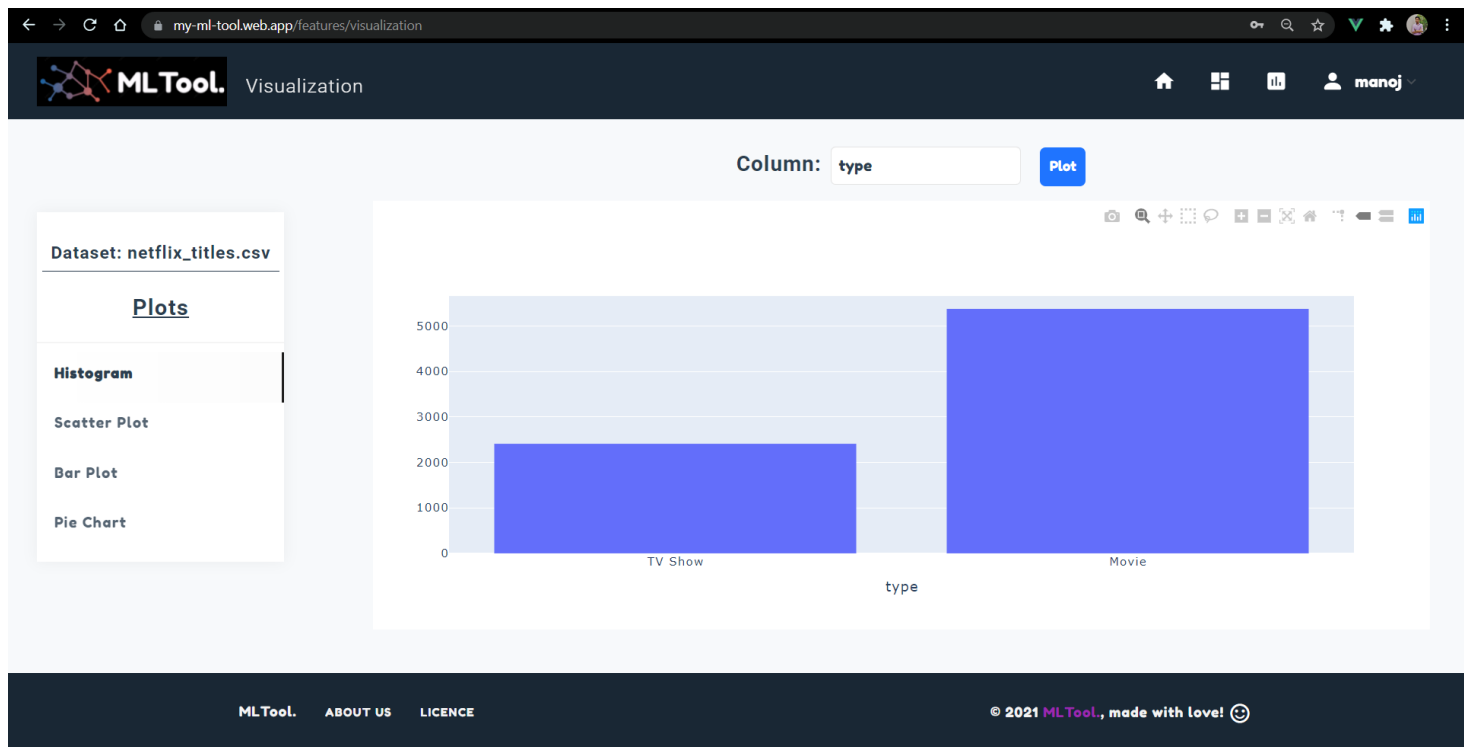
Please select

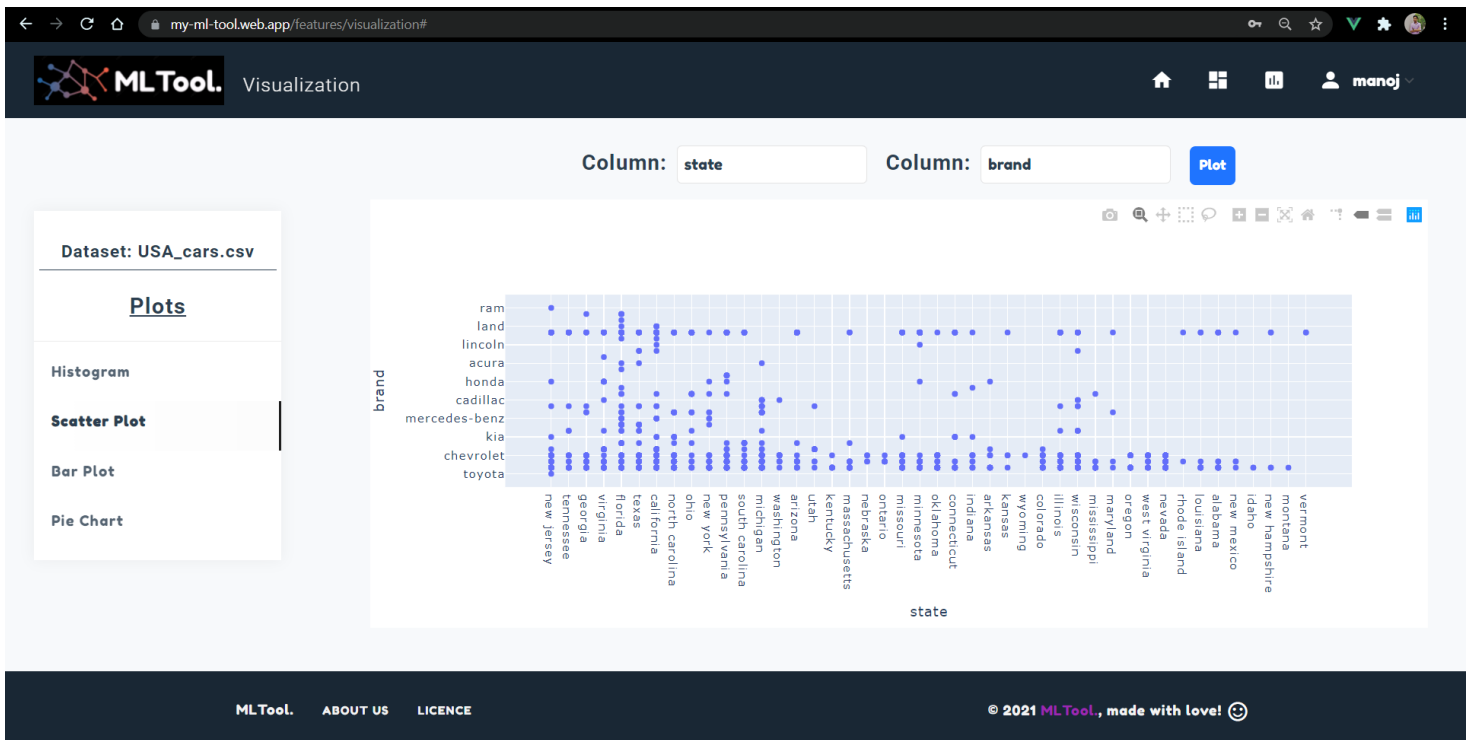
Normalise/Scale dataset using:

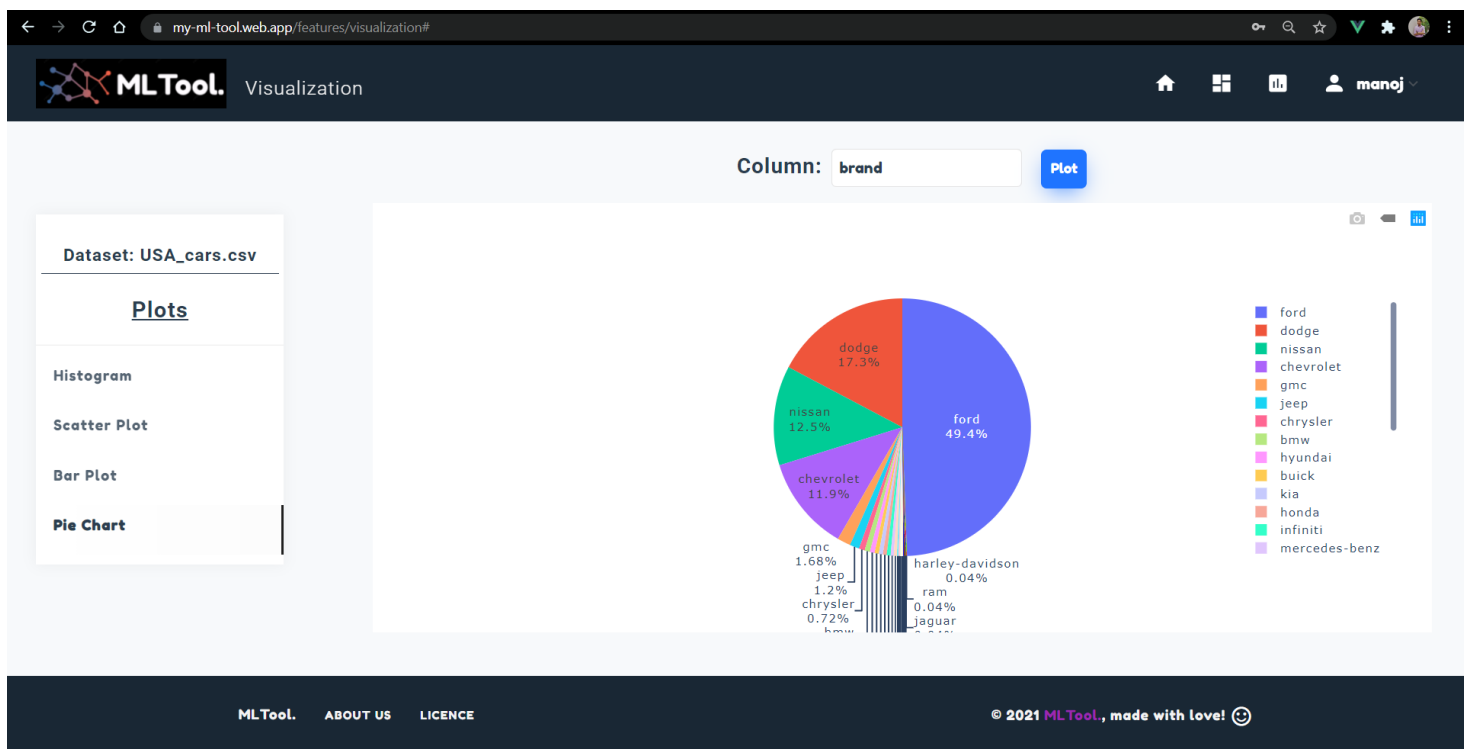
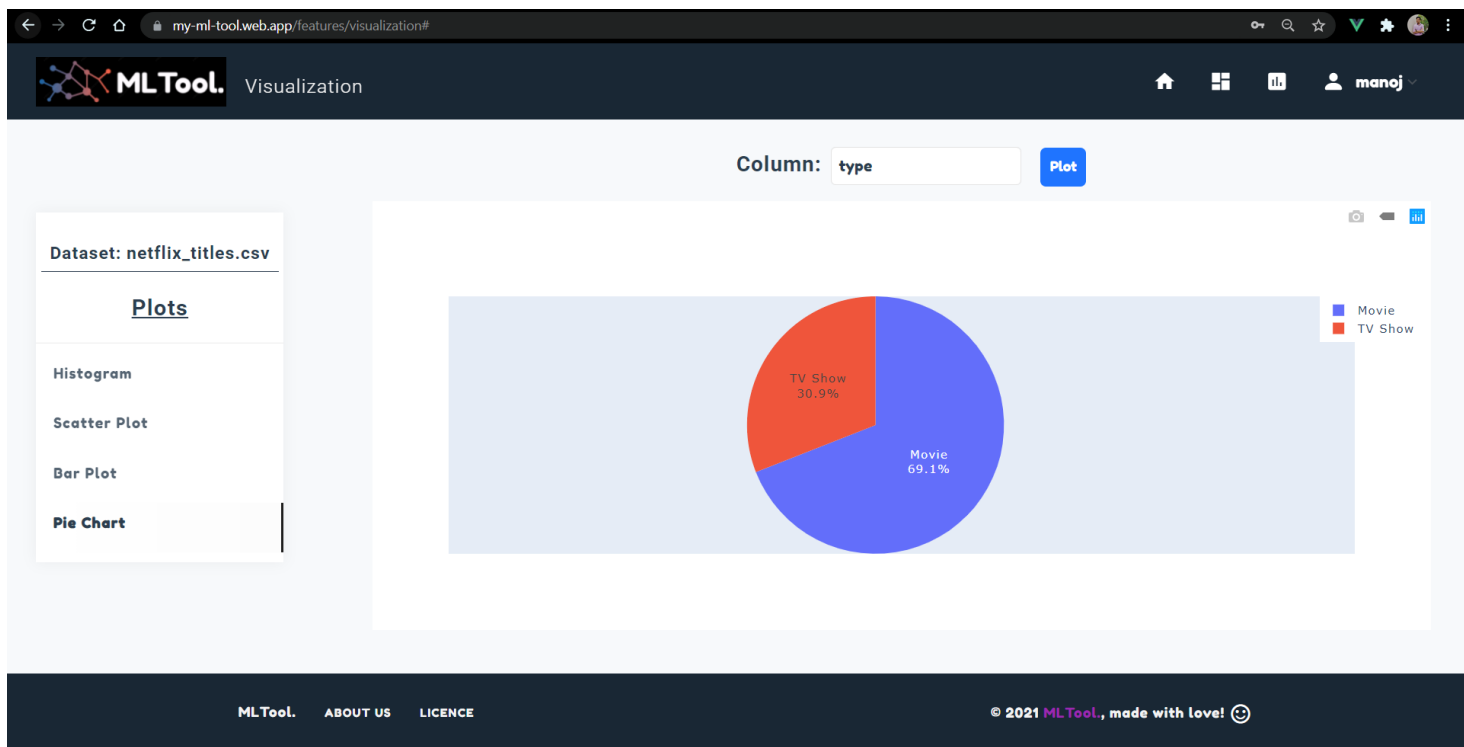
Optional

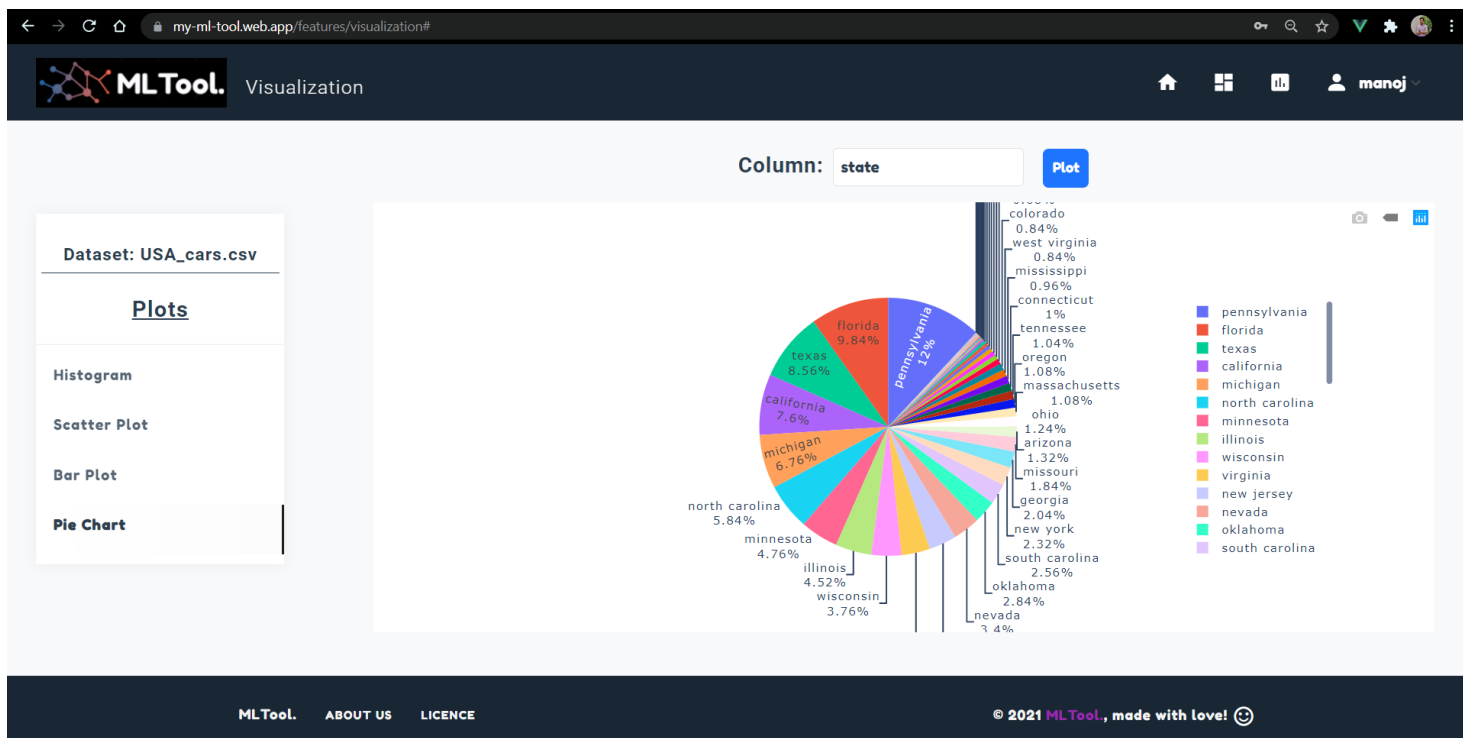
Pre-Process Data

Dataset Visualization:









ML Model Training:

my-ml-tool.web.app/features/models

MLTool. Models Home

Select columns to train the model for: netflix_titles.csv

Target Column: type

Select all columns ☐

- ☒ country
- ☒ date_added
- ☒ release_year
- ☒ rating
- ☒ duration
- ☒ listed_in

MLTool. ABOUT US LICENCE © 2021 MLTool., made with love! 😊

Choose the model to train the data:

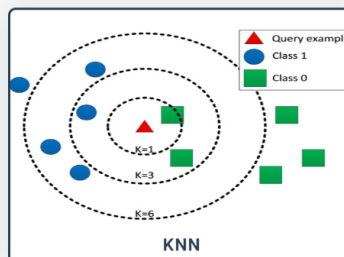
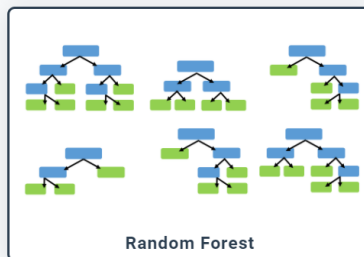
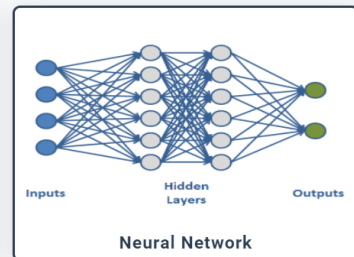
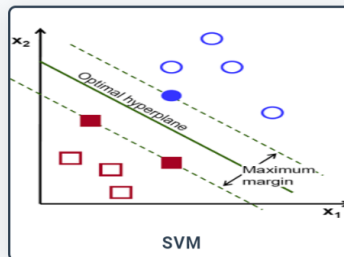
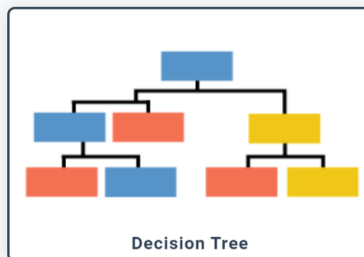


Diagram illustrating the Naive Bayes formula:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Labels and arrows in the diagram:

- Likelihood** points to $P(x|c)$.
- Class Prior Probability** points to $P(c)$.
- Posterior Probability** points to $P(c|x)$.
- Predictor Prior Probability** points to $P(x)$.

Below the formula, the joint probability expression is shown:

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

Naive Bayes

Parameters for Decision Tree

Not sure what parameters to use? [Get Best Parameters](#)

Note: Below auto-filled parameters are the default values set by sklearn.

Criterion:

Gini

Splitter:

Best

Maximum Depth of the tree: (integer)
(If None, then nodes are expanded until all leaves are pure)

None

Minimum number of samples required to split an internal node:

2

Minimum number of samples required to be a leaf node:

1

Maximum number of leaf nodes (integer):
(If None then unlimited number of leaf nodes.)

None

Train Model

← → ↺ 🏠 my-ml-tool.web.app/features/models 🔑 🔍 ☆ 🟢 ⚙️ 👤 ⋮

Parameters for KNN

Not sure what parameters to use? [Get Best Parameters](#)

Note: Below auto-filled parameters are the default values set by sklearn.

Number of neighbors (k):

Weights:

Algorithm:

Leaf size:

Power parameter:

[Train Model](#)

MLTool. ABOUT US LICENCE


© 2021 MLTool., made with love! 😊

ML Model Report:

← → ↺ 🏠 my-ml-tool.web.app/features/models/model-report 🔑 🔍 ☆ 🟢 ⚙️ 👤 ⋮

ML Model Report

Model Chosen: **Decision Tree**



Accuracy: 0.955963

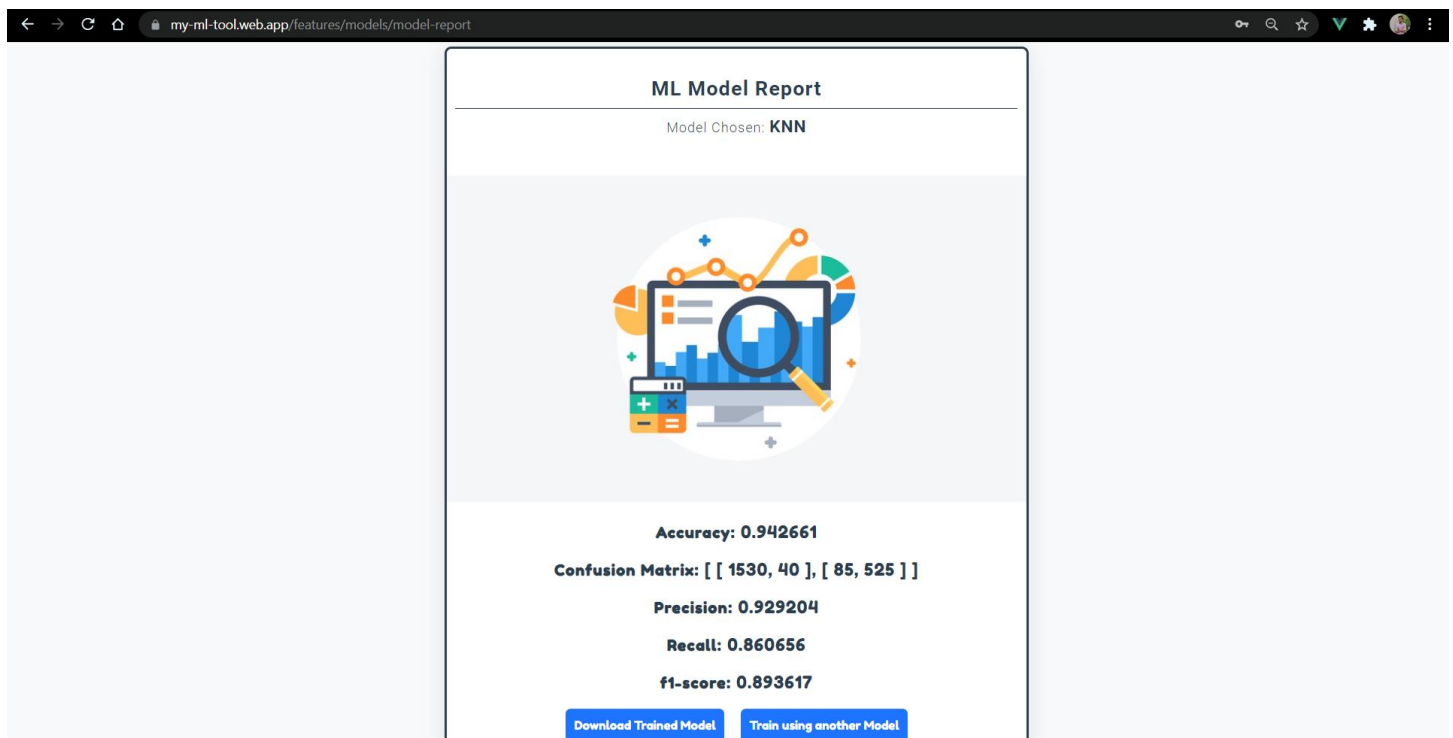
Confusion Matrix: [[1558, 12], [84, 526]]

Precision: 0.977695

Recall: 0.862295

f1-score: 0.916376

[Download Trained Model](#) [Train using another Model](#)



Watch the full demo of the application here:

<https://www.dropbox.com/s/h11afe4xx6xwqtu/ML%20Tool%20Demo%20Video.mp4?dl=0>