Technology Review

SciKit-Learn & Tensorflow

Krisha Mehta
Anant Rajeev
Leena Elamrawy

Outline

- Project Overview
- Use Case
- Libraries
- Advantages of the libraries
- Drawbacks of the libraries

Project Overview

- Renewable energy and alternative fuels are being more utilized in the world nowadays, the most notable of which is the advancement of hybrid, electric, and hydrogen powered vehicles
    - According to iea.org, electric vehicle registrations have increased 40% each year over the past two years
    - Car companies like Tesla, Toyota, Nissan, and Honda are experimenting with hybrid and off-gas models at a higher rate
- With these hybrid, full electric, and full hydrogen powered vehicles becoming more popular nationwide, more charging stations in the country are needed.
- It's clear that alternative fuels are the future. Can the U.S. infrastructure keep up with it and can the U.S. supply those users with the resources it needs to keep their alternative-fuel vehicles running for a long time

Use Case

- Our project has 2 main goals that we want to achieve
  - 1) Classification Tool
    - Users are able to interact with the interface to input details about an electric vehicle that they want and their personal profile, and our backend clustering model will, based on the data we have, output some vehicles that are popular and they might like with some statistics to back it up
  - 2) Visualization Tool
    - Through an interactive visualization tool, users can see many statistics about electric vehicles, including charging stations across the United States, individual vehicle attributes, and individual vehicle vs. vehicle comparisons

Scikit-learn and Tensorflow

SCIKIT-LEARN

- General purpose machine learning library (traditional), prefers users to process data by themselves such as selecting features, compressing dimensions, etc.
- Low degree of freedom, but increases the efficiency of the model and reduces the difficulty of batching and standardization (through the use of pipelines).
- Suitable for small and medium-sized projects that have a small amount of data and can be completed on the CPU and low hardware requirements.

TENSORFLOW

- Deep learning library, it automatically extracts valid features from the data and does not need to be done manually.
- High degree of freedom, the algorithms can be implemented by ourselves
- Suitable for projects having a large data, low data processing requirements and ultimately require GPU-accelerated operations.

Advantages of Scikit-learn and Tensorflow

SCIKIT-LEARN

- Very well documented
- Covers most ML tasks
- Simple and efficient tools for machine learning and statistical modeling
- Build on NumPy, SciPy and matplotlib
- Open source and commercially usable

TENSORFLOW

- Extensible and powerful library
- Much better performance and customizability than Sklearn
- Open source platform for machine learning
- Robust ML production anywhere and powerful experimentation for research

Drawbacks of Scikit-learn and Tensorflow

SCIKIT-LEARN

- Doesn't work well for unsupervised learning applications
- Inability to reasonably perform AutoML

TENSORFLOW

- Requires a solid understanding of ML
- Learning curve is very steep
- Using TensorFlow requires a separate installation
- Does not have a strong visualization framework
- Tensor API is not very user friendly

REFERENCES

- Image (https://media.ed.edmunds-media.com/kia/niro-ev/2019/fd/2019_kia_niro-ev_f34_fd_204194_1600.jpg)
- https://scikit-learn.org/stable/
- https://www.tensorflow.org/overview
- https://www.iea.org/reports/global-ev-outlook-2020