

# AI and ML Landscape and its Python Libraries

Dr. S V Viraktamath\*, Veena Goudru\*\*, Vismay S Nargund\*\*\*, Venkatesh S Trimalle\*\*\*\*

\*(Dept. of Electronics and Communication, SDM College of Engineering and Technology, and Dharwad, India

Email:svvmath@gmail.com)

\*\* (Dept. of Electronics and Communication, SDM College of Engineering and Technology, and Dharwad, India

Email :veenagoudruveen@gmail.com)

\*\*\* (Dept. of Electronics and Communication, SDM College of Engineering and Technology, and Dharwad, India

Email :vismay.nargund@gmail.com)

\*\*\*\* (Dept. of Electronics and Communication, SDM College of Engineering and Technology, and Dharwad, India

Email :venkst9945@gmail.com)

\*\*\*\*\*

## Abstract:

The rapid evolution of artificial intelligence (AI) and Machine Learning (ML) has revolutionized various industries. Python owing to its simplicity, versatility and a vast ecosystem of libraries has become a powerful tool that facilitates research and development in these domains. This review paper provides an analysis of why python is preferred language for AI and ML delving into its key attributes and widespread adoption. After which fundamentals of Artificial Intelligence and Machine Learning are discussed emphasizing their significance in various domains and role of python in it. The paper also highlights a selection of python libraries developed using AI and ML. Some of the prominent libraries discussed include Tensor flow, Pytorch, scikit-learn and Keras. Furthermore, some applications like facial recognition, data clustering and drug-target interaction are discussed to showcase how python libraries using AI and ML are helpful in the physical world.

**Keywords —Machine Learning, Artificial Intelligence, Python Libraries, Tensor Flow, SciKit-Learn**

\*\*\*\*\*

## I. INTRODUCTION

In the ever-changing landscape of modern technology, the use of artificial intelligence (AI) and machine learning with python has emerged as a transformative force, changing the way we use technology. Python being a versatile and widely adopted technology among both experts and beginners has now evolved to become a bedrock for AI and ML applications thanks to its extensive ecosystem of libraries and framework that facilitate the development of AI and ML solutions. Thispaper delves into the relationship shared between python and the realms of AI and ML, exploring its implications and far-reaching applications in diverse domains [1].As we transverse through the sections of this paper, we will embark in a comprehensive exploration, beginning with an introduction to python, this will act as a foundation for our understanding of AI and ML in the following section where we explore its key components and principles.Our focus however will remain on python as we

navigate through the myriad of libraries dedicated for AI and ML. From popular libraries like TensorFlow, Pytorch, Keras to recent ones like Albi, Tianshou and River. We will dissect these libraries, highlighting their unique capabilities and applications that empower developers and data scientists to unlock the fun potential of AI and ML.

Beyond theory and frameworks, we venture into real world applications that leverage the prowess of AI and ML with python. From healthcare and finance to image recognition and natural language processing demonstrating how AI and ML algorithms enhance efficiency, accuracy and innovation. Therefore, this paper serves as a comprehensive guide to the fusion of python with AI and ML shedding light on the core libraries, their functionalities and their applications, hoping to instil a deep appreciation for the dynamic duo of python and AIML, and inspire further innovation in this ever-evolving field.

## II. PYTHON

Python is a high-level programming language known for its versatility, simplicity and readability. It was created by Goidovon Rossum with the vision of creating a language that prioritizes readability and ease of use. Discovered in 1991, it is a powerful, procedural, object-oriented, and functional programming language. The logo of python is shown in figure-1. Due to its wide applications, it has steadily grown to gain immense popularity in fields like development and data science [2] as shown in figure-2. Python also has a thriving community that contributes to its growth and development. The Python Package Index (PyPI) hosts a plethora of third-party libraries and frameworks, always updating Python’s capabilities.



Figure1. Logo of Python [3]

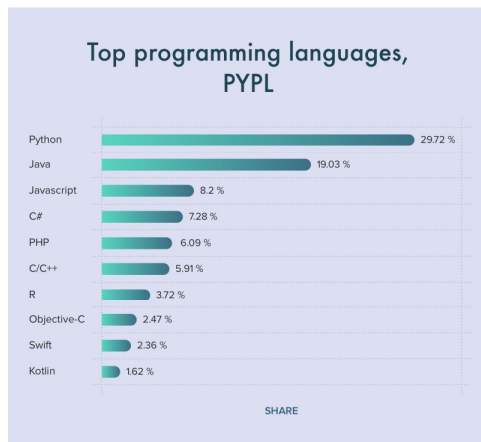


Figure2. Popularity of Python as preferred language to code in[4].

It is also an ideal choice for both beginners and experienced programmers thanks to its clear and concise syntax as evident in figure-5.

```

1 int main( )
2 {
3     float num1, num2, sum;
4     printf ("please input two numbers\n");
5     scanf ("%f %f", &num1, &num2);
6     sum=num1 + num2;
7     printf (" the sum is %f \n", sum);
8     return 0;
9 }
    
```

Figure3. C program to add two numbers [5].

```

1 import java.util.Scanner;
2 public class HelloWorld {
3     public static void main( String args[] ) {
4         Scanner sc= new Scanner(system.in);
5         system.out.println( "please input the first number:");
6         float num1=Float.parseFloat(sc.nextLine( ));
7         System.out.println( "please input the second number:");
8         float num2=Float.parseFloat(sc.nextLine( ));
9         float sum=num1+num2;
10        System.out.println("the sum is"+sum);
11    }
12 }
    
```

Figure4. Java code to add two numbers[5].

```

1 num1 = input(" please input the first number: ")
2 num2 = input(" please input the second number: ")
3 sum = float( num1) + float(num2)
4 print("sum is", sum)
    
```

Figure5. Python code to add two numbers[5].

### A. WORKING

Python codes are written in a text editor or an integrated development environment (IDE) and is typically saved in files with a “.py” extension. It is an interpreted language where the code is compiled into bytecode, which is a lower-level representation of the code and is stored with a “.py” extension. This compilation happens automatically when you run a python script. When the python script or program is run the interpreter translates the byte code into machine code and executes it line by line [6]. The working of python is shown in figure-6.

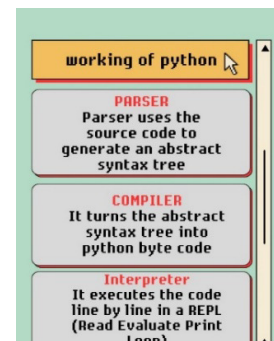


Figure6. Working of Python

### B. CHARACTERISTICS of PYTHON

- Python is easier to learn with a straight forward syntax and being much the same as English language.
- It doesn't require the use of semi-colon or curly-bracket as the indentation itself defines the code block.
- It is a dynamically typed language i.e., there is no need to specify variable types explicitly. The interpreter determines the types of the variable at runtime suppose an integer value of 5 is assigned to x, then just `x=5` is used to declare the variable and not `int x=25`.
- It also boasts of a vast range of libraries for various fields such as scripting, web development and machine learning. Django, flask and pyramids are the popular frameworks for python. While Tensor flow, Numpy and Pytorch are some popular choices for machine learning.
- Python with another functionality also provides GUI programming facilities allowing for code to be written in such a way that user interface for any application can be created easily by just importing libraries like Tkinter, Flexx, Data, CEF Python etc.
- Python is a multi-paradigm programming language and therefore supports object-oriented style where codes are encapsulated within objects. Therefore, everything written in python is in the form of classes and objects [7]

### C. TOP TECHNOLOGICAL COMPANIES USING PYTHON

- **Google:** Python is one of the primary languages used by Google's data division. It is heavily utilized in many apps and is given priority support. Google employees are among the major contributors to python and they actively use, support and promote the language and

often sponsors PyCon and other python conferences [8]. The logo of Google is shown in figure-7.



Figure7. Logo of Google[9].

- **YouTube:** YouTube serves as one example which uses python for managing website designs managing videos, viewing videos, gaining access to canonical data and numerous other functions [8]. The logo of YouTube is shown in figure-8.



Figure8. Logo of YouTube[10].

- **Netflix:** Python is the primary language used by Netflix. Python is used across the whole content lifecycle, from selecting which content to find to managing the content delivery network (CDN) that provides the finished movie to more than 148 million users [8]. The logo of Netflix is shown in figure-9.



Figure9.Logo of Netflix[11].

- **Facebook/Instagram:** Python is Facebook's and Instagram's primary programming language, particularly in AI and data science. For instance, python is used by the Facebook oculus team to create virtual reality (VR) hardware and software [8].The logo of Instagram is shown in figure-10.



Figure10.Logo of Instagram[12].

- **NASA:** A significant portion of NASA's software is open source and uses python for data and image processing related to spatialdata. You can check out many pythonprograms at [code.nasa.gov](http://code.nasa.gov) [8].The logo of NASA is shown in figure-11



Figure11.Logo of NASA [13].

- **PayPal:** The whole IT architecture of PayPal is managed with python. The primary programming language used by their brain free SDK for express checkout integration and other APIs is python. Also, a sizeable quantity of python code and kits are accessible on theirwebsite [8].The logo of paypal is shown in figure-12



Figure12.Logo of PayPal[14].

### III. ARTIFICIAL INTELLIGENCE

Intelligence is defined as the ability to understand and learn things while Artificial Intelligence is defined as a subfield of computer science emphasizing on designing computer programs and machines capable of performing tasks which mimic human intelligent behavior such as recognition, judgement, thinking, reasoning, learning etc. To performing tasks that are natural to humans such as natural language understanding, decision making, speech comprehension, image recognition [15].

#### A. Categories of Machine Learning

- **Narrow AI:**This type of AI is designed to perform specifications or solve particular problems. It operates within a limited domain and doesn't possess general intelligence. For example, Siri Chatbot. Python is easier to learn with a straight forward syntax and being much the same as English language.
- **General AI:** General AI, also known as Artificial General Intelligence, is a theoretical concept of AI that possess human-like cognitive abilities and can understand learn and apply knowledge in a wide-range of tasks. AGI is not yet realised and remains as object of ongoing research.
- **Machine Language:** It is a subfield of AI as shown in figure-13 that focuses on developing algorithms and models that enable machines to learn from data and make predictions or decision trees based on data and is widely used in recognition and recommendation system.

- **Deep Learning:** Deep learning is a subset of machine language that uses artificial neural network inspired by the structure of human brain to solve complex problems. It has been particularly successful in tasks such as image and speech recognition.

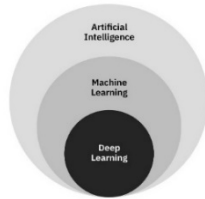


Figure13. ML as a subset of AI [16].

### B. CHARACTERISTICS of AI

- **AI is data driven:** It relies on data for learning, decision making and improving its performance in various tasks. The more relevant and high-quality data the AI system has access to the better it can perform.
- **AI is Self-learning:** AI systems can learn and improve their performance without explicit human intervention. These AI systems typically employ machine learning techniques to adapt and refine over time.
- **AI can respond in real time:** Artificial Intelligence can think and react to circumstances in real-time by using neural network to deepen data analysis.
- **AI reaches very high accuracy:** Deep neural networks enable AI to operate with extremely high precision. This capability has helped with the AI application’s ability to diagnose conditions like cancer[17].

### IV. MACHINE LANGAUGE(ML)

Machine Learning is a subset of artificial intelligencethat focuses on the development of algorithms and models that enable computers to learn from and make predictions or decisions based on data. In machine learning, the system doesn’t rely on explicit programming instructions instead it learns from patterns and information contained in the data[18].Figure-14 illustrates the different categories of machine learning.

### A. CATEGORIESof MACHINE LEARNING

- **Supervised learning:** It is a type of machine learning where an algorithm learns from labelled training data to make prediction or decisions without human intervention. In supervised learning the algorithm is “supervised”, in the sense that it learns to map input data to a known labelled output. The primary goal is to get the machine to learn mapping input features to a target output.
- **Unsupervised learning:** In this category of machine learning, an algorithm is trained on data without explicit supervision. Instead, the algorithm explores the data’s inherent structure and patterns on its own.
- **Reinforcement learning:** It is a type of machine learning where an agent learns to make sequence of decisions by interacting with an environment. The agent’s objective is to maximise a cumulative reward signal overtime by selecting actions that lead to favourable outcomes. Reinforcement learning is inspired by behavioural psychology and is used to solve problems that involved decision-making and sequential actions[19].

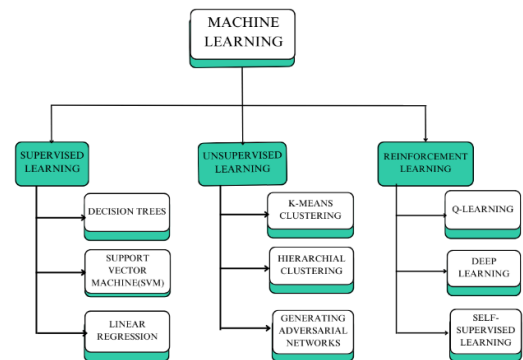


Figure14. Different categories of Machine Learning Algorithms

### V. PYTHON LIBRARIES

Libraries are essentially a bunch of pre-written code and functions that programmers can build upon to perform common tasks or access specific functionality without having to write the code from scratch. This significantly increases the development process and reduce the likelihood of errors. Many libraries follow

established standards and best practices, which helps maintain consistency as well as ensures the code is reliable and well-tested [20].

#### A. POPULAR LIBRARIES for AI and ML PROJECTS

**TensorFlow:** One of the popular and free open-source libraries for machine learning in python. It mainly focuses on training and inferring deep neural networks. The multi-dimensional arrays known as Tensors, which allow it to perform multiple operations on a particular input. Also, Tensor Board a feature of Tensor flow helps you to visualise graphs and learn the models. Another powerful tool is Graph dashboard, which give a quick view of the model's design and structure. Further the APIs of Tensor flow is hierarchically arranged and low-level API's act as a foundation on which high-level APIs are built[21].

**SciKit-Learn:** Learning Python libraries offers a variety of supervised and unsupervised learning algorithms through a compatible interface. It uses a variety a technique, including K-means density based spatial clustering of application with noise (DBSCAN), support vector machines, decision trees, naive Bayes, random forests and regression with support vector machine to come up with effective algorithm. Moreover, it communicates with Numpy and SciPy, two python numerical libraries for numerical related applications. Similarly, the SciKit-image library provides a collection of image processing techniques, that covers segmentation, analysis, filtering, feature recognition in images, colour space, manipulation and morphology [22].

**Keras:** It is a powerful open-source library for building and turning neural network which can be used effectively for various machine learning tasks. Known for its user-friendly high-level API, Keras allows you to create models by stacking various layers together and can also run on top of several deep learning frameworks including Tensor flows, Theano and CNTK

recently Keras has been tightly integrated with Tensor flow and is referred as "tf. keras"[23].

**Numpy:** A Math python library called Numpy is a library for processing massive multi-dimensional arrays and matrices, with the aid of an extensive set of advanced mathematical functions. It also facilitates effective and efficient calculation. It has proven to be useful for scientific computation with python in machine learning and artificial intelligence. In a similar when it provides a comprehensive N-dimensional array interference and a linear algebra function that are orders of magnitude quicker and more memory efficient than other python lists, and therefore is a crucial library for simulation and machine learning [24].

**SciPy:** Based on python NumPy extension. SciPy is an open-source python library which is used to solve math, science, engineering and technical problems. Allowing for manipulation as well as visualization of data by making use of wide range of python command. It also provides specialized tools for array computing [25].

#### B. NEW PYTHON LIBRARIES for AI and ML

**Albi:** It is a python library designed for explaining machine learning models, which provides a framework for model interpretation and understanding. It particularly focuses on generating explanations for individual predictions and therefore the library provides various techniques and tools to help users gain insights into how their machine learning model takes decision and makes prediction. Albi is not tied to a specific type of machine learning model and can be used to explain various machine learning models like deep neural networks, decision trees, random forest and more. It provides wide range of explanation methods which include features like feature attribution, counter factual generation and adversarial attacks and helps uses understand the most influential features as input

components of a given prediction, which can be vital for understanding the inner workings of a model and making it more interpretable. Another important feature is that the Alibi supports the generation of counterfactual explanations which is especially useful for understanding the conditions under which a model might change its output. It also provides tools to visualize and present explanations in a clear and intuitive manner. Further Alibi can be integrated with popular machine learning frameworks like tensor flow and Pytorch increasing its compatibility and making it easier to use [26].

**River:** It is a python library designed for machine learning with specific focus on streaming data with machine learning framework that provides tools and algorithms for efficiently handling data streams. This makes it suitable for real-time and online learning scenarios where data continuously flows in and updates need to be made on the fly. It is designed as an online machine learning where models are updated continuously as new data arrives. It allows you to train models on data streams and adapt to changing patterns over time which means it focuses on incremental learning where the model is updated continuously as new data arrives. This makes it memory-efficient and well-suited for streaming data. Further it supports active learning where the model itself decides which data points to label using techniques like model evaluation. River is highly flexible and customisable which is very valuable when adapting machine learning models to specific use cases and domains. Further River has an active community and continuously being developed and improved [27].

**Tianshou:** It is an open-source deep reinforcement learning (DRL) library developed by the Tencent AI lab. It is designed to provide a highly modular and flexible framework for building and experimenting with various DRL algorithms written in python and

built on top of Pytorch. Tianshou is structured in such a way that it allows users to easily mix and match components to create custom reinforcement learning pipelines. It does so by providing a range of pre-implemented components such as different algorithms, models and exploration strategies that can be combined and extended to suit various tasks. Tianshou supports both single-agent and multi-agent reinforcement learning and uses popular DRL algorithms like DQN, A2C, PPO and more. It can be used for tasks like continuous control, discrete action games. Tianshou is often used by researchers and practitioners to conduct experiments, develop and test new reinforcement learning algorithms and adapt existing algorithms to specific problem domains. Its modular design and extensive documentation make it a valuable tool for deep reinforcement learning [28].

**Pandas:** One of the most potent python libraries for data manipulation, analysis and cleaning is called Pandas. It facilitates the import, analysis and visualization of data by working with rational and labelled data. It is a library for data analysis that offers an extensive range of tools for working with high-level data structures. Pandas come with built-in features for grouping, filtering and merging data in accordance with time-series functionality [29].

**PYML:** 'PYML' as associated with the Penn machine learning project is a python library designed for machine learning research and experimentation. It offers a collection of machine learning algorithms for classification, regression and more with specified focus on support vector machine. PYML also includes utilities for data preprocessing and feature engineering offering a range of evaluation metrics to assess the performance of machine learning models, such as accuracy, precision, recall, F1 score. It may also include Hyperparameter optimization for further tuning and optimization of machine learning model [30].

## VI. APPLICATIONS

- **Face-detection and recognition:** It is the most popular application of artificial intelligence technology. It is the process of identifying a person by mapping facial features by using various methods as shown in the figure-15. In this project python programming language has been used along with Open CV library and machine learning library for reliable real-time face detection and recognition. The computer vision is a python library having built-in features and advanced capabilities such as face tracking, face detection, face recognition, Kalman filtering and many multi-platform frameworks that supports windows, Linux and mac operating systems enhancing the applicability across multiple platforms[20].

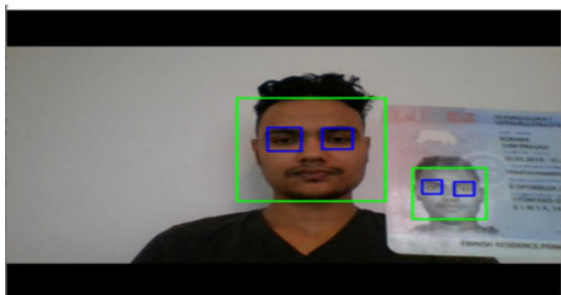


Figure15. Face detection using Open CV[20].

- **Clustering-** It is a type of unsupervised approach that is frequently utilized in various fields for statistical data analysis clustering. Typically involving breaking up a collection of observations into smaller groups or clusters, so that observations within the same clusters are similar to each other and observations outside of the same cluster are distinct. Thus, grouping data accordingly by similarity and diversity. In this project to generate two-dimensional data set that contained four blobs, as shown in figure-16 Thus, successfully demonstrating method of clustering data using python and its libraries[15].

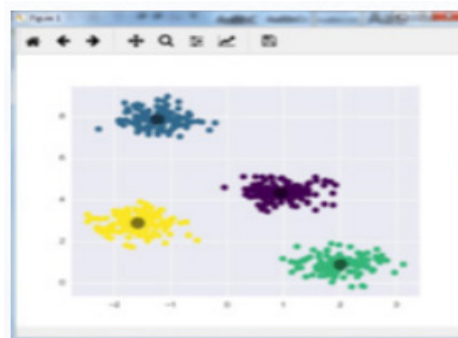


Figure16. Data Clustering[15].

- **Drug-target predictions:** The task of predicting the interactions between drugs and targets an example of which is shown in figure-17 plays a key role in the process of drug discovery. Therefore, a python package and PyDPI was developed. It is based on Random Forest and integrates Proteochemometrics, bioinformatics and chemo genomics for PTI prediction using a framework involving the selection of molecular features and predefined dictionaries for classification[31]

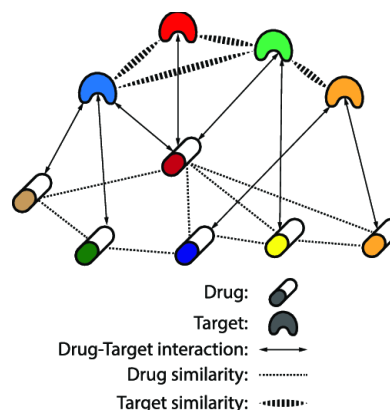


Figure17. Drug-Target interactions[31].

## VII. CONCLUSION

This review paper has illuminated the pivotal role of python libraries in the realm of artificial intelligence and machine learning. The adoption of python as the language of choice for AI and ML development has catalysed innovation and advancements in various industries. This review paper has emphasized the significance of key python libraries such as Tensor flow, Pytorch, a kit-learn and Keras has revealed the diversity of tools available for AI and ML. These libraries



give the ease and flexibility of implementing complex models, experimenting with machine learning algorithms and address a diverse set of problems. The paper also delved into applications of python libraries after demonstrating how implementation of AI and ML using python has various real-world efficacies using facial recognition, data clustering and Drug-target interaction as examples. In essence, this review paper serves as a comprehensive guide for researchers, practitioners and enthusiasts in the field of AI and ML it provides a understanding of python libraries essential for AI and ML development. While also shedding light on their practical applications As the AI and ML domain continues to develop, the significance of python libraries is also going to increase and the applications arising out of it are sure to further push the boundaries of what is possible.

#### REFERENCES

- [1] Ku. Chhaya A. Khanzode, Dr. Ravindra D. Sarode, "ADVANTAGES AND DISADVANTAGES OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING: A LITERATURE REVIEW", International Journal of Library & Information Science (IJLIS), Volume 9, Issue 1, January-April 2020
- [2] Masoud Nosrati "Python: An appropriate language for real world Programming" World Applied Programming, Vol (1), No (2), June 2011
- [3] <https://pluspng.com/python-logo-png-10014.html>
- [4] <https://www.cleveroad.com/blog/programming-languages-ranking>
- [5] Yongcheng Wu "The Course Construction and Reform of Python Programming in the Background of Artificial Intelligence", 5th International Conference on Social Science and Higher Education.
- [6] A.L.SayethSaabith, T. Vinohraj, M.M. Fareez "Popular python libraries and their application domains" International Journal of Advance Engineering and Research Development Volume 7, Issue 11, November-2020
- [7] Nimit Thaker, Abhilash Shukla, "Python as Multi Paradigm Programming Language", International Journal of Computer Applications Volume 177 – No. 31, January 2020
- [8] A. Rawat, "A Review on Python Programming", IJRESM, vol. 3, no. 12, pp. 8–11, Dec. 2020.
- [9] <https://pluspng.com/google-photos-logo-png-5947.html>
- [10] <https://www.brandsoftheworld.com/logo/youtube-10?original=1>
- [11] <https://www.pngarts.com/explore/15581>
- [12] <https://pluspng.com/instagram-png-978.html>
- [13] <https://www.pngall.com/nasa-png/>
- [14] <https://www.pngall.com/paypal/download/52545>
- [15] Aleksandar Miljković, Slobodan Nedeljković, Milan Čabarkapa, "APPLICATION OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN PYTHON PROGRAM LANGUAGE", International scientific conference serba, 18-19 November 2020
- [16] Machine learning (ibm.com)
- [17] R. A. Corlett, "Features of artificial intelligence languages and their environments", Software Engineering Journal July 1986
- [18] Pinky Sodhia, Naman Awasthi, Vishal Sharma, "Introduction to Machine Learning and Its Basic Application in Python", Proceedings of 10th International Conference on Digital Strategies for Organizational Success, 31 Jan 2019
- [19] Jie Zhang, Shu Liang, Xi Liu Zhou, "Application and Analysis of image recognition technology based on Artificial Intelligence-machine learning algorithm as an example", 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL)
- [20] Shiv Bohara, "Artificial Intelligence and Machine Learning", Centria University of Applied Sciences Information Technology, June 2020.
- [21] Samira Gholizadeh, "Top Popular Python Libraries in Research", Journal of Robotics and Automation Research, February 2022
- [22] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot and Edouard Duchesnay, "Scikit-learn: Machine Learning in Python", Journal of Machine Learning Research 12 (2011) 2825-2830
- [23] Bahzad Taha Chicho, Amira Bibo Sallow, "A Comprehensive Survey of Deep Learning Models Based on Keras Framework", University Tun Hussein Onn Malaysia Publisher's Office, JOURNAL OF SOFT COMPUTING AND DATA MINING VOL. 2 NO. 2 (2021) 49-62
- [24] S. van der Walt, S. C. Colbert and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," in Computing in Science & Engineering, vol. 13, no. 2, pp. 22-30, March-April 2011
- [25] Dr. V. Hanuman Kumar, "Python Libraries, Development Frameworks and Algorithms for Machine Learning Applications", International Journal of Engineering Research & Technology (IJERT), Vol. 7 Issue 04, April-2018.
- [26] Janis Klaise, Arnaud Van Looveren, Giovanni Vacanti, "Alibi Explain: Algorithms for Explaining Machine Learning Models", Journal of Machine Learning Research 22 (2021) 1-7
- [27] Jacob Montiel, Max Halford, Saulo Martiello Mastelini, Geoffrey Bolmier, Raphael Sourty, Robin Vaysse, Adil Zouitine, Heitor Murilo Gomes, Jesse Read, Talel Abdesslem, Albert Bifet, "River: machine learning for streaming data in Python", Journal of Machine Learning Research 22 (2021) 1-8
- [28] Jiayi Weng, Huayuan Chen, Dong Yan, Kaichao You, Alexis Duburq, Minghao Zhang, Yisu, Hang Su, Jun Zhu, "Tianshou: A Highly Modularized Deep Reinforcement Learning Library.", Journal of Machine Learning Research 23 (2022)
- [29] Wes McKinney, "pandas: a Foundational Python Library for Data Analysis and Statistics", Unpublished 2011
- [30] Davide Albanese, Roberto Visintainer, Stefano Merler, Samantha Riccadonna, Giuseppe Jurman, Cesare Furlanello, "mlpy: Machine Learning Python", arXiv: 1202.6548v2[cs.MS] 1 Mar 2012
- [31] Maryam Bagherian, Elyas Sabeti, Kai Wang, Maureen A. Sartor, Zaneta Nikolovska-Coleska and Kayvan Najarian, "Machine learning approaches and databases for prediction of drug-target interaction: survey paper", Briefings in Bioinformatics, 22(1), 2021, 247–269