**PROFESSIONAL INDUSTRIAL TRAINING**

**/ INTERNSHIP REPORT**

**ON**

**“Human Voice Tone Emotion Detection”**

**AT**

**Solitaire Infosys Pvt. Ltd.**

**Sahibzada Ajit Singh Nagar, Punjab 160055.**

**AN INDUSTRY INTERNSHIP REPORT SUBMITTED**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS**

**FOR THE AWARD OF DEGREE OF**

**BACHELOR OF ENGINEERING**

**In**

**Computer Science & Engineering**

**SUBMITTED BY**

Ananya Gupta

Roll Number:2020A1R139

**A picture containing text, clipart

Description automatically generated**

SUBMITTED TO

Department of Computer Science & Engineering

Model Institute of Engineering and Technology (Autonomous)

Jammu, India

2022

**CANDIDATES’ DECLARATION**

I, **Ishav Verma, 2020A1R160,** hereby declare that the work which is being presented in the Industry Internship Report entitled, “**Human Voice Tone Emotion Detection**” in partial fulfillment of requirement for the award of degree of B.E. Computer Science and Engineering and submitted in the Department of Computer Science and Engineering, Model Institute of Engineering and Technology (Autonomous), Jammu is an authentic record of my own work carried by me at “Solitaire Infosys, Punjab” under the supervision and mentorship of **Gagan Sawhney,** HR, Solitaire Infosys, Punjab and Miss Shafalika Vijyal, Assistant Professor, Dept. of CSE, MIET respectively. The matter presented in this report has not been submitted in this or any other University / Institute for the award of B.E. Degree.

*Signature of the Student*  *Dated*: 25-11-2022

Ananya Gupta

**2020A1R139**

**INTERNSHIP CERTIFICATE**

****

**Department** **of Computer Science & Engineering**

**Model Institute of Engineering and Technology (Autonomous)**

**Kot Bhalwal, Jammu, India**

***(NAAC “A” Grade Accredited)***

**Ref. No.: Date:**

**CERTIFICATE**

Certified that this Industry Internship Report entitled **“Human Voice Tone Emotion Detection”** is the bonafide work of **“Ananya Gupta, 2020A1R139, of 5th Semester, Computer Science & Engineering, Model Institute of Engineering and Technology (Autonomous), Jammu”,** who carried out the Industry Internship at “Solitaire Infosys Pvt. Ltd, Punjab” work under my mentorship during July-2022 to August-2022.

**Miss Shafalika Vijyal**

**Mentor-Internal Supervisor**

**Assistant Professor**

**Department of CSE, MIET**

*This is to certify that the above statement is correct to the best of my knowledge.*

**(Name)**

**Designation**

**Head of the Department**

**Department Name, MIET**

**ACKNOWLEDGEMENTS**

This Summer internship opportunity was a great chance for learning and professional development. I am grateful for having a chance to meet so many wonderful people and professionals who led me though this internship period.

It is my pleasant duty to pay my heartfelt gratitude to Mr. Gagan Sawhney, HR, Solitaire Infosys who have guided me through the course of this Internship.

I must record my deep sense of gratitude to Prof. (Dr.) Ankur Gupta (Director, MIET) and Prof. (Dr.) Ashok Kumar (Dean Academics) for their guidance, constant inspiration and encouragement, and for their keen involvement throughout the course of present work.

Gratitude and thanks although mean a very small thing to convey my thanks to my parents who have always given me a parental source of love, motivation and strength right from the journey of my life.

Bearing in mind previous I am using this opportunity to express my deepest gratitude and special thanks to the teachers who in spite of being extraordinarily busy with their duties, took time out to hear, guide and keep me on the correct path and allowing me to carry out my project at their esteemed organization and extending during the training.

I perceive this opportunity as a big milestone in my career development. I will strive to use gained skills and knowledge in the best possible way, and I will continue to work on their improvement, in order to attain desired career objectives. Hope to continue cooperation with all of you in the future.

I express my sincere gratitude to Solitaire Infosys Pvt. Ltd. and Model Institute of Engineering and Technology (Autonomous), Jammu for giving me the opportunity.

**Ananya Gupta**

**2020A1R139**

**SELF EVALUATION**

I am a 3rd year B.E. undergraduate student pursuing Computer Science and Engineering at Model Institute of Engineering and Technology, Jammu. I recently completed an internship with Solitaire Infosys Pvt. Ltd Sahibzada Ajit Singh Nagar, Mohali, Punjab as a Machine Learning Intern.

There I learned about the Machine Learning and its applications in day-to-day lives. Also, I learned about the Python Programming concepts in a very easy and efficient manner.

I was also provided with multiple assessments during my internship, which I always completed on time with full dedication and zeal. I still experienced a learning curve due to this being my first exposure to this kind of work. By the end of my internship, however, I felt comfortable in completing my assigned tasks and even received reviews from team leaders expressing their opinions about my work.

I developed great communication skills with people and this helped me to be a good team member. when difficult situations occurred in meeting a deadline or solving a problem. Teamwork is valuable to me because I welcome coworker insights into these types of challenges.

I totally understand the importance of regular practice and learning conceptual theories while being a CS student. And due to this internship opportunity, I got the chance to learn the topics not only theoretically but practically too. I got a firmer grasp on the coding part and learned a lot of new concepts.

While working as a Machine Learning Intern at Solitaire Infosys Pvt. Ltd, I gained a newer kind of experience which is surely going to help me for my future endeavors.

Ananya Gupta

2020A1R139

**ABSTRACT**

As advancement in healthcare is going on so also on Mental Health. According to an survey, “Mental illnesses are common in the United States. Nearly one in five U.S. adults live with a mental illness (52.9 million in 2020)”. And studies have proven that human voice can be used as a parameter to predict his mood and so on after evaluating the feedback further we can also get to know about the mental health of the individual. So, from here we got our problem statement of “Human Voice Tone Emotion Detection”. Under this project we are creating a Deep Learning based model to analyze the same. We are using Convolutional Neural Networks for same. Where we are using Normalization, Convo1D, Max Pooling & Dense layers to build a Sequential Model with TensorFlow & Keras with a total of 12,85,390 parameters out of which 12,84,622 trainable parameters resulting in precise breakdown of input file to get the most accurate prediction of it.

**Contents**

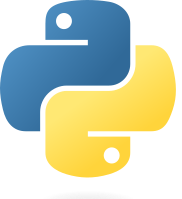
|  |  |  |
| --- | --- | --- |
|  | Candidates’ Declaration | i |
|  | Internship Certificate | ii |
|  | Certificate | iii |
|  | Acknowledgement | iv |
|  | Self-Evaluation | v |
|  | Abstract | vi |
|  | Contents | vii |
|  | List of Figures | ix |
|  | Abbreviations Used | x |
| **Chapter 1 Python** | | **1-3** |
| 1.1 | Introduction to Python | 1 |
| 1.2 | History of Python | 1 |
| 1.3 | Development in Python | 1 |
| 1.4 | Features of Python | 2 |
| 1.5 | Use of Python | 2 |
| **Chapter 2 Artificial Intelligence** | | **4-8** |
| 2.1 | About Artificial Intelligence | 4 |
| 2.2 | History of Artificial Intelligence | 5 |
| 2.3 | Types of Artificial Intelligence | 7 |
| 2.4 | Applications of Artificial Intelligence | 7 |
| **Chapter 3 Machine Learning** | | **9-14** |
| 3.1 | About Machine Learning | 9 |
| 3.2 | Difference - Human Learning & Machine Learning | 9 |
| 3.3 | Difference - Rule Based Approach & Machine Learning | 10 |
| 3.4 | Problems solved using Machine Learning | 10 |
| 3.5 | Types of Machine Learning | 11 |
| 3.6 | Procedure of Machine Learning | 14 |
| **Chapter 4 Deep Learning** | | **15-17** |
| 4.1 | About Deep Learning | 15 |
| 4.2 | Neural Networks | 15 |
| 4.3 | Difference - Machine Learning & Deep Learning | 16 |
| **Chapter 5 Modules & Libraries** | | **18-19** |
| 5.1 | Python Modules | 18 |
| 5.2 | Python Libraries | 18 |
| **Chapter 6 Project Description** | | **20-23** |
| 6.1 | Problem Statement | 20 |
| 6.2 | Workflow of Project | 20 |
| 6.3 | Tech Stack Used | 20 |
| 6.4 | Dataset Used | 21 |
| 6.5 | Significance of important code segments | 22 |
| 6.6 | Accuracy & Loss | 22 |
| **References** | | **24** |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **Figure Number** | **Figure Title** | **Page Number** |
| 2.1.1 | Applications of AI | 4 |
| 2.1.2 | Concepts related to AI | 4 |
| 2.2.1 | History of AI | 5 |
| 2.4.1 | Application fields of AI | 7 |
| 3.2.1 | Human Learning | 9 |
| 3.2.2 | Machine Learning | 9 |
| 3.3.1 | Rule Based Approach | 10 |
| 3.3.2 | Machine Learning Approach | 10 |
| 3.5.1 | Supervised Machine Learning | 11 |
| 3.5.2 | Unsupervised Machine Learning | 12 |
| 3.5.3 | Unsupervised Machine Learning - Clustering | 12 |
| 3.5.4 | Semi-supervised Machine Learning | 13 |
| 3.5.5 | Reinforcement Learning. | 13 |
| 3.6.1 | Machine Learning Process | 14 |
| 4.2.1 | Neural Network Representation | 15 |
| 4.3.1 | Deep Learning Advantage | 16 |
| 6.2.1 | Workflow | 20 |
| 6.5.1 | Model Definition | 22 |
| 6.5.2 | Model Compilation | 22 |

**ABBREVIATIONS USED**

|  |  |
| --- | --- |
| AI | Artificial Intelligence |
| ANN | Artificial Neural Network |
| CNN | Convolutional Neural Network |
| DL | Deep Learning |
| ML | Machine Learning |
| NN | Neural Networks |
| RNN | Recurrent Neural Network |

**Chapter 1**

**Python**

*Python is the gift that keeps on giving.*

*The more you understand Python, the more you can do in the 21st Century. As simple as that.*

* 1. **Introduction to Python**

Python is a widely-used, interpreted, object-oriented, and high-level programming language with dynamic semantics, used for general-purpose programming. It’s everywhere, and people use numerous Python-powered devices on a daily basis, whether they realize it or not.

* 1. **History of Python**

Python was created by Guido van Rossum, and first released on February 20, 1991. While you may know the python as a large snake, the name of the Python programming language comes from an old BBC television comedy sketch series called Monty Python’s Flying Circus.

One of the amazing features of Python is the fact that it is actually one person’s work. Usually, new programming languages are developed and published by large companies employing lots of professionals, and due to copyright rules, it is very hard to name any of the people involved in the project. Python is an exception.

Of course, Guido van Rossum did not develop and evolve all the Python components himself. The speed with which Python has spread around the world is a result of the continuous work of thousands (very often anonymous) programmers, testers, users (many of them aren’t IT specialists) and enthusiasts, but it must be said that the very first idea (the seed from which Python sprouted) came to one head – Guido’s.

* 1. **Development in Python**

Python is maintained by the Python Software Foundation, a non-profit membership organization and a community devoted to developing, improving, expanding, and popularizing the Python language and its environment.

* 1. **Features of Python**

Python is omnipresent, and people use numerous Python-powered devices on a daily basis, whether they realize it or not. There are billions of lines of code written in Python, which means almost unlimited opportunities for code reuse and learning from well-crafted examples. What’s more, there is a large and very active Python community, always happy to help.

There are also a couple of factors that make Python great for learning:

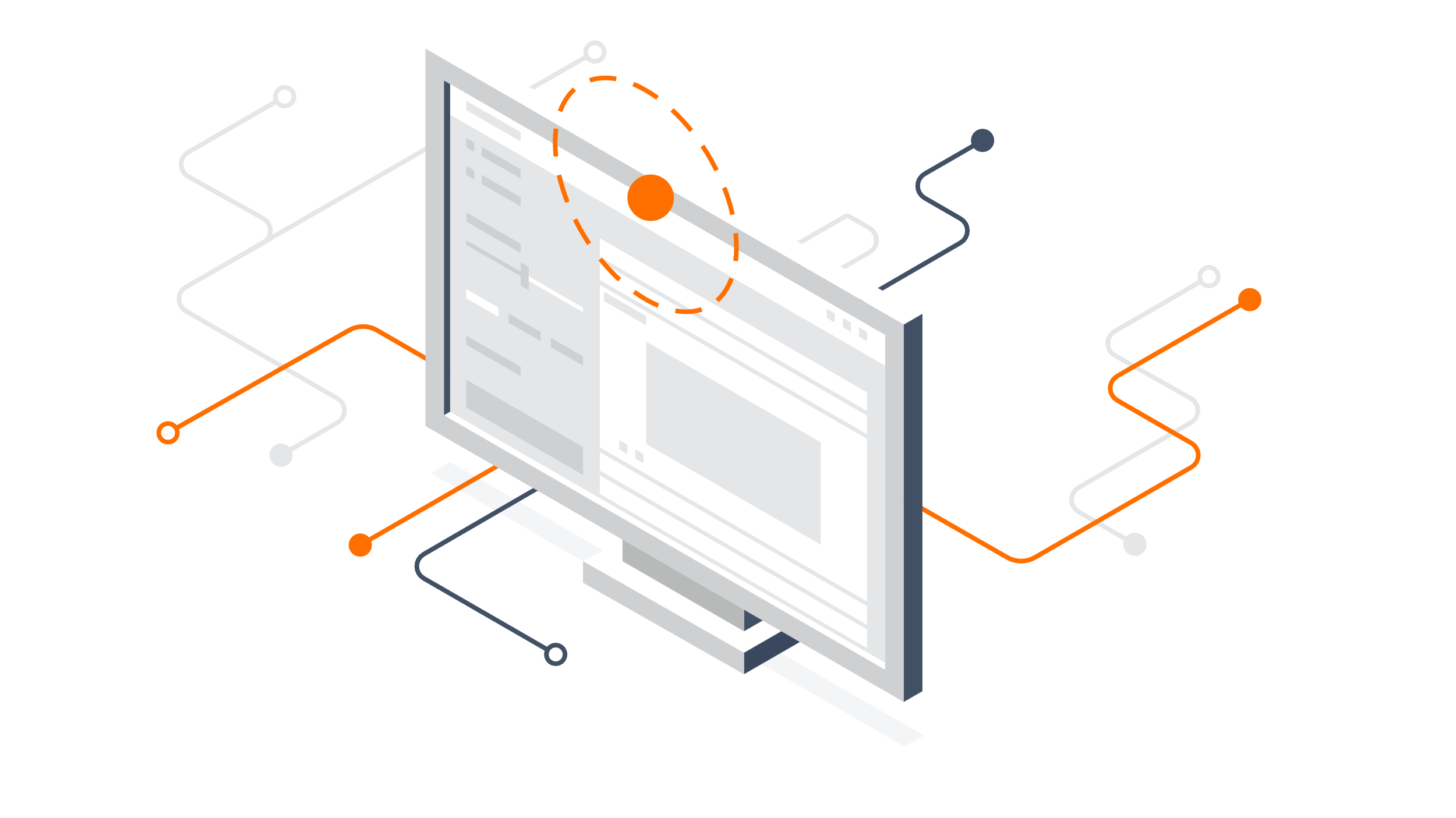
* It is easy to learn – the time needed to learn Python is shorter than for many other languages; this means that it’s possible to start the actual programming faster;
* It is easy to use for writing new software – it’s often possible to write code faster when using Python;
* It is easy to obtain, install and deploy – Python is free, open and multiplatform; not all languages can boast that.
  1. **Use of Python**

Programming skills prepare you for careers in almost any industry, and are required if you want to continue to more advanced and higher-paying software development and engineering roles. Python is the programming language that opens more doors than any other. With a solid knowledge of Python, you can work in a multitude of jobs and a multitude of industries. And the more you understand Python, the more you can do in the 21st Century. Even if you don’t need it for work, you will find it useful to know.

Many developing tools are implemented in Python. More and more everyday use applications are being written in Python. Lots of scientists have abandoned expensive proprietary tools and switched to Python. Lots of IT project testers have started using Python to carry out repeatable test procedures. The list is long.

Python is a great choice for:

* Web and Internet development (e.g., Django and Pyramid frameworks, Flask and Bottle micro-frameworks)
* Scientific and numeric computing (e.g., SciPy – a collection of packages for the purposes of mathematics, science, and engineering; Ipython – an interactive shell that features editing and recording of work sessions)
* Education (it’s a brilliant language for teaching programming!)
* Desktop GUIs (e.g., wxWidgets, Kivy, Qt)
* Software Development (build control, management, and testing – Scons, Buildbot, Apache Gump, Roundup, Trac)
* Business applications (ERP and e-commerce systems – Odoo, Tryton)
* Games (e.g., Battlefield series, Sid Meier’s Civilization IV…), websites and services (e.g., Dropbox, UBER, Pinterest, BuzzFeed...)

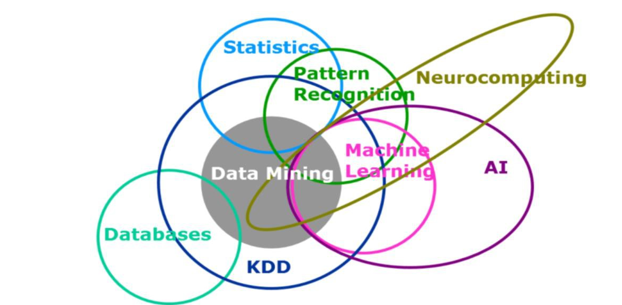
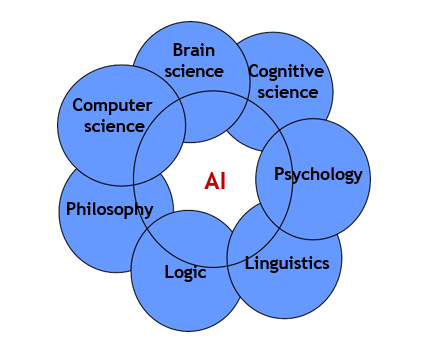
**Chapter 2**

**Artificial Intelligence**

*Mankind is welcoming the fourth industrial revolution represented by intelligent technology. New technologies such as AI integrated into all aspects of human society; driving change in global macro trends, such as sustainable social development and economic growth. New kinetic energy, smart city upgrading, industrial digital transformation, consumer experience, etc*

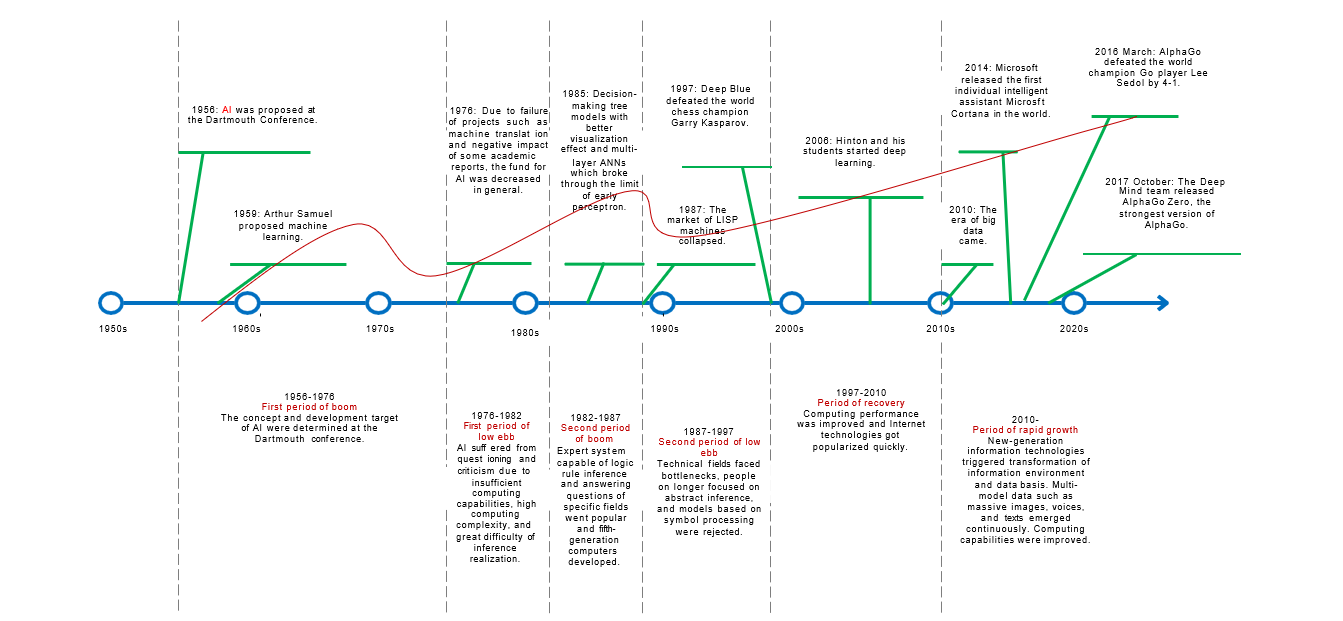
* 1. **About Artificial Intelligence**

Artificial Intelligence (AI) is a new technical science that studies and develops theories, methods, techniques, and application systems for simulating and extending human intelligence. In 1956, the concept of AI was first proposed by John McCarthy, who defined the subject as "science and engineering of making intelligent machines, especially intelligent computer program". AI is concerned with making machines work in an intelligent way, similar to the way that the human mind works. At present, AI has become an interdisciplinary course that involves various fields.



**Fig 2.1.1:** Application of AI **Fig 2.1.2:** Concepts related to AI

* 1. **History of Artificial Intelligence**

****

**Fig 2.2.1:** History of AI.

From 1957 to 1974, AI flourished. Computers could store more information and became faster, cheaper, and more accessible. Machine learning algorithms also improved and people got better at knowing which algorithm to apply to their problem. Early demonstrations such as Newell and Simon’s General Problem Solver and Joseph Weizenbaum’s ELIZA showed promise toward the goals of problem solving and the interpretation of spoken language respectively. These successes, as well as the advocacy of leading researchers (namely the attendees of the DSRPAI) convinced government agencies such as the Defense Advanced Research Projects Agency (DARPA) to fund AI research at several institutions. The government was particularly interested in a machine that could transcribe and translate spoken language as well as high throughput data processing. Optimism was high and expectations were even higher. In 1970 Marvin Minsky told Life Magazine, “from three to eight years we will have a machine with the general intelligence of an average human being.” However, while the basic proof of principle was there, there was still a long way to go before the end goals of natural language processing, abstract thinking, and self-recognition could be achieved.

In the 1980’s, AI was reignited by two sources: an expansion of the algorithmic toolkit, and a boost of funds. John Hopfield and David Rumelhart popularized “deep learning” techniques which allowed computers to learn using experience. On the other hand Edward Feigenbaum introduced expert systems which mimicked the decision making process of a human expert. The program would ask an expert in a field how to respond in a given situation, and once this was learned for virtually every situation, non-experts could receive advice from that program. Expert systems were widely used in industries. The Japanese government heavily funded expert systems and other AI related endeavors as part of their Fifth Generation Computer Project (FGCP). From 1982-1990, they invested $400 million dollars with the goals of revolutionizing computer processing, implementing logic programming, and improving artificial intelligence. Unfortunately, most of the ambitious goals were not met. However, it could be argued that the indirect effects of the FGCP inspired a talented young generation of engineers and scientists. Regardless, funding of the FGCP ceased, and AI fell out of the limelight.

Ironically, in the absence of government funding and public hype, AI thrived. During the 1990s and 2000s, many of the landmark goals of artificial intelligence had been achieved. In 1997, reigning world chess champion and grand master Gary Kasparov was defeated by IBM’s Deep Blue, a chess playing computer program. This highly publicized match was the first time a reigning world chess champion loss to a computer and served as a huge step towards an artificially intelligent decision-making program. In the same year, speech recognition software, developed by Dragon Systems, was implemented on Windows. This was another great step forward but in the direction of the spoken language interpretation endeavor. It seemed that there wasn’t a problem machines couldn’t handle. Even human emotion was fair game as evidenced by Kismet, a robot developed by Cynthia Breazeal that could recognize and display emotions.

One could imagine interacting with an expert system in a fluid conversation, or having a conversation in two different languages being translated in real time. We can also expect to see driverless cars on the road in the next twenty years (and that is conservative). In the long term, the goal is general intelligence, that is a machine that surpasses human cognitive abilities in all tasks. This is along the lines of the sentient robot we are used to seeing in movies. Even if the capability is there, the ethical questions would serve as a strong barrier against fruition. When that time comes (but better even before the time comes), we will need to have a serious conversation about machine policy and ethics (ironically both fundamentally human subjects), but for now, we’ll allow AI to steadily improve and run amok in society.

* 1. **Types of Artificial Intelligence**

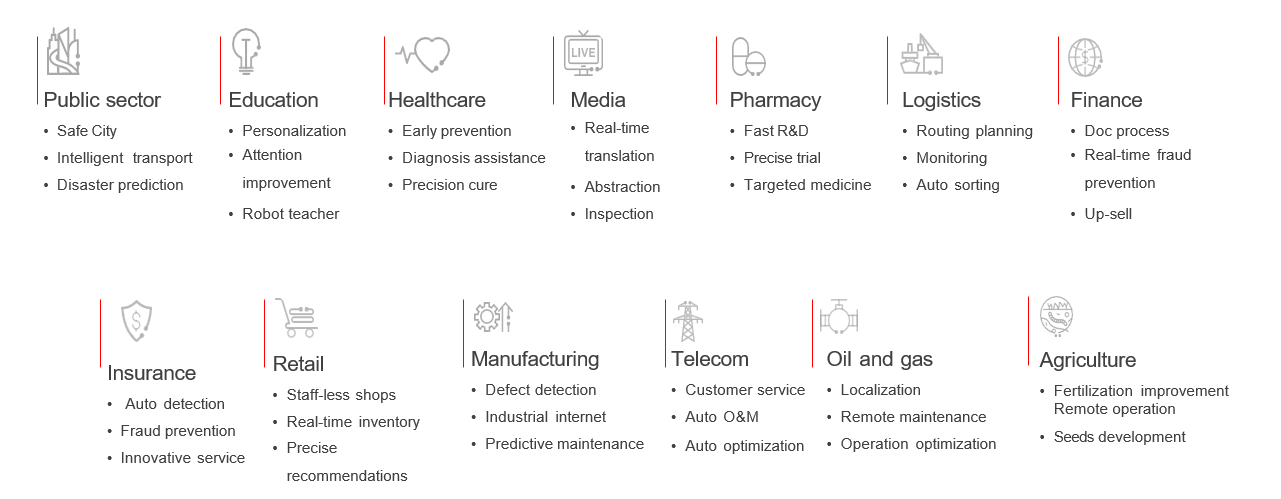
Generally, AI can be classified as two distinct types. They are as:

* *Strong AI:* The strong AI view holds that it is possible to create intelligent machines that can really reason and solve problems. Such machines are considered to be conscious and self-aware, can independently think about problems and work out optimal solutions to problems, have their own system of values and world views, and have all the same instincts as living things, such as survival and security needs. It can be regarded as a new civilization in a certain sense.
* *Weak AI:* The weak AI view holds that intelligent machines cannot really reason and solve problems. These machines only look intelligent, but do not have real intelligence or self-awareness.
  1. **Applications of Artificial Intelligence**

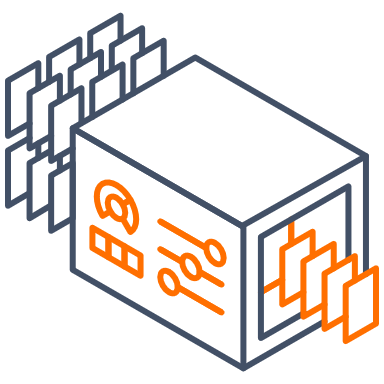
At present, application directions of AI technologies mainly include:

* *Computer vision:* a science of how to make computers "see".
* *Speech processing:* a general term for various processing technologies used to research the voicing process, statistical features of speech signals, speech recognition, machine-based speech synthesis, and speech perception.
* *Natural language processing (NLP):* a subject that use computer technologies to understand and use natural language

Also, these technologies can be categorized as sub-fields according to their use in that field.



**Fig 2.4.1:** Application Fields of AI.

**Chapter 3**

**Machine Learning**

*Machine learning is a core research field of AI, and it is also a necessary knowledge for deep learning.*

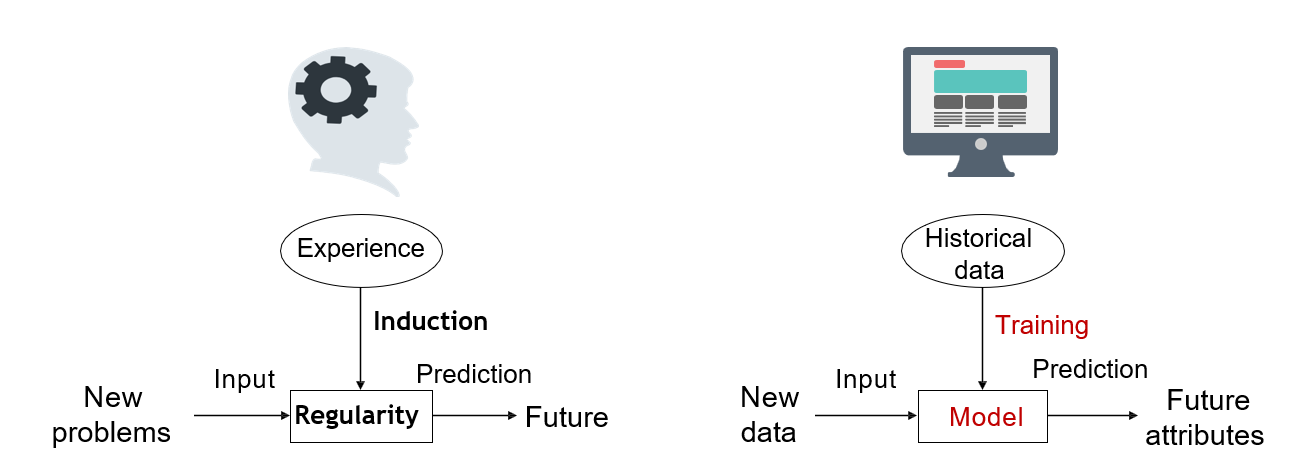
**3.1 About Machine Learning**

Machine learning (including deep learning) is a study of learning algorithms. A computer program is said to learn from experience 𝐸 with respect to some class of tasks 𝑇 and performance measure 𝑃 if its performance at tasks in 𝑇, as measured by 𝑃, improves with experience 𝐸.

**3.2 Difference - Human Learning & Machine Learning**

Humans acquire knowledge through experience either directly or shared by others. Machines acquire knowledge through experience shared in the form of past data.

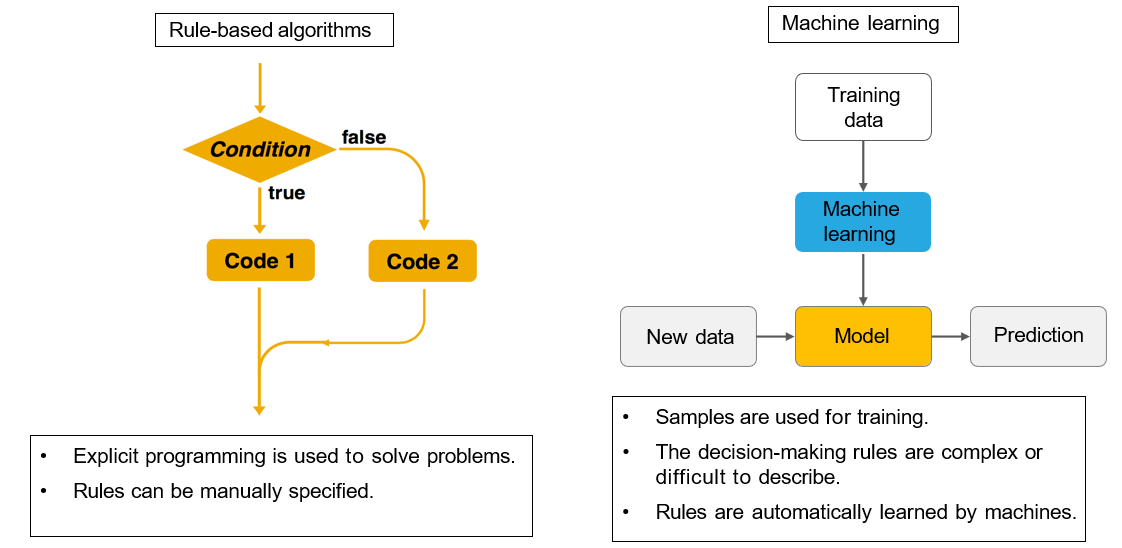
Skill is a manifestation of intelligence possessed by humans. And intelligence is the ability to apply knowledge. Human intelligence sustains, but his knowledge fades as new technologies emerge. Humans without knowledge in particular subjects can apply their intelligence to solve problems in new domains. But machines can solve new problems only if their intelligence has been updated with retraining on data acquired from the changed scenarios. This is a fundamental difference between human intelligence and machine intelligence.



**Fig 3.2.1:** Human Learning **Fig 3.2.2:** Machine Learning

**3.3 Difference – Rule Based Approach & Machine Learning**

In contrast with Rule-Based Approach Machine Learning is more advantageous in aspects of variable input values. For saying if about Rule Based approach, we have to manually define the set of rules on which the data is processed and output is produced. But on the other hand in case of Machine Learning we use data as for training our model & in accordance to data provided the model itself defines a boundary of rules that are complex but more effective as compared to Rule Based.



**Fig 3.3.1:** Rule Based Approach **Fig 3.3.2:** Machine Learning Approach

**3.4 Problems Solved using Machine Learning**

Machine learning can deal with many types of tasks. The following describes the most typical and common types of tasks.

* *Classification:*

A computer program needs to specify which of the k categories some input belongs to. To accomplish this task, learning algorithms usually output a function 𝑓: 𝑅𝑛 → (1,2, … , 𝑘). For example, the image classification algorithm in computer vision is developed to handle classification tasks.

* *Regression:*

For this type of task, a computer program predicts the output for the given input. Learning algorithms typically output a function 𝑓: 𝑅𝑛 → 𝑅. An example of this task type is to predict the claim amount of an insured person (to set the insurance premium) or predict the security price.

* *Clustering:*

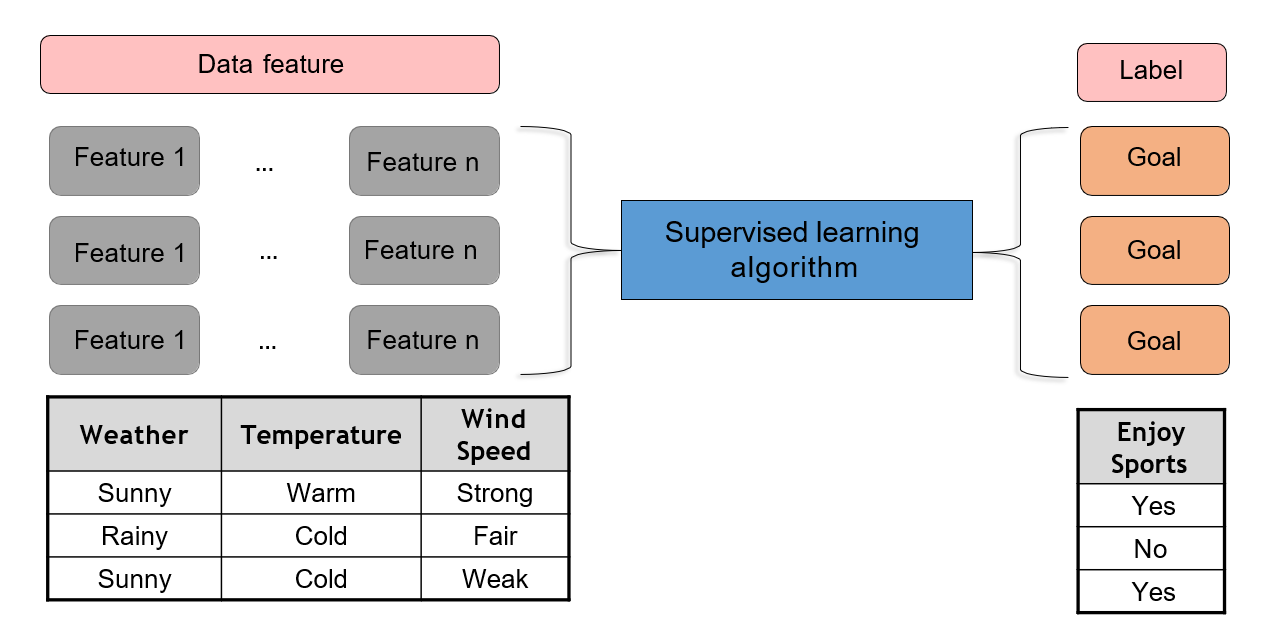
A large amount of data from an unlabelled dataset is divided into multiple categories according to internal similarity of the data. Data in the same category is more similar than that in different categories. This feature can be used in scenarios such as image retrieval and user profile management.

**3.5 Types of Machine Learning**

In general, Machine Learning is classified in four types. They can be highlighted as:

* *Supervised learning:*

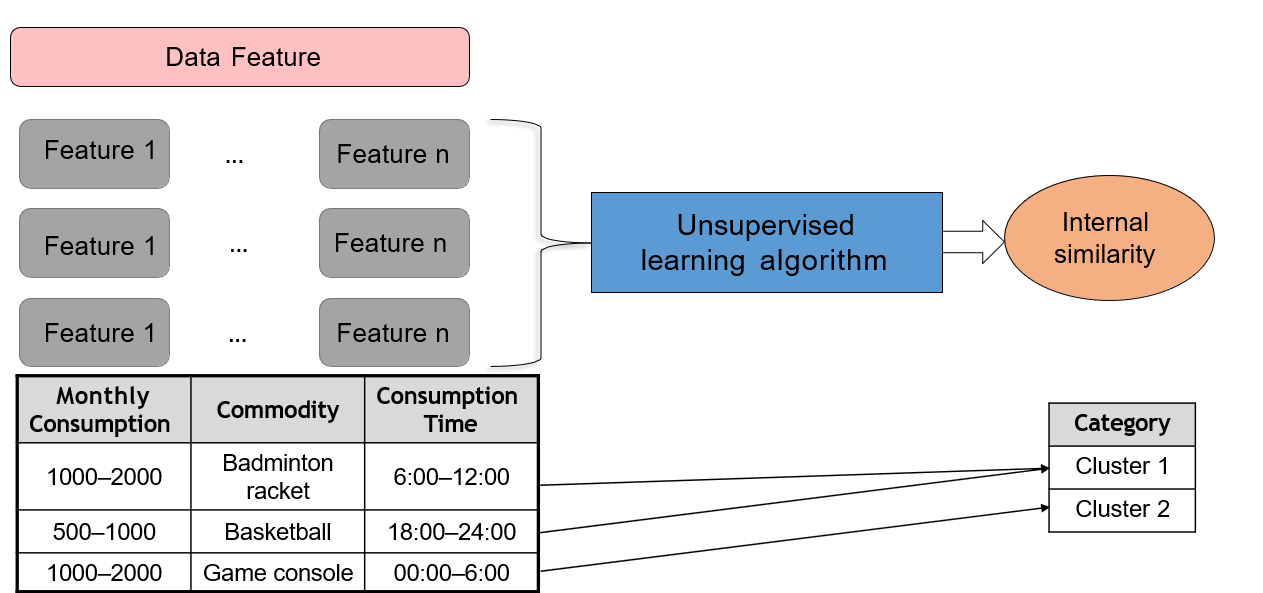
Obtain an optimal model with required performance through training and learning based on the samples of known categories. Then, use the model to map all inputs to outputs and check the output for the purpose of classifying unknown data.

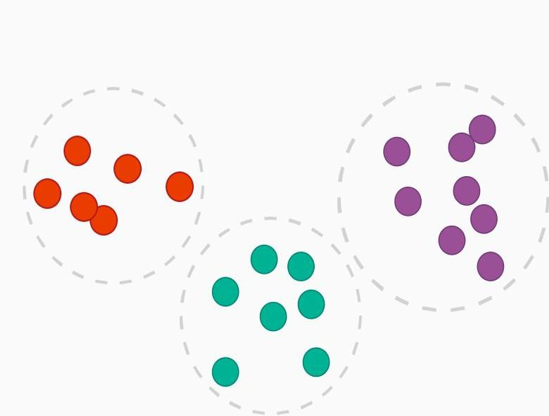


**Fig 3.5.1:** Supervised Machine Learning.

* *Unsupervised learning:*

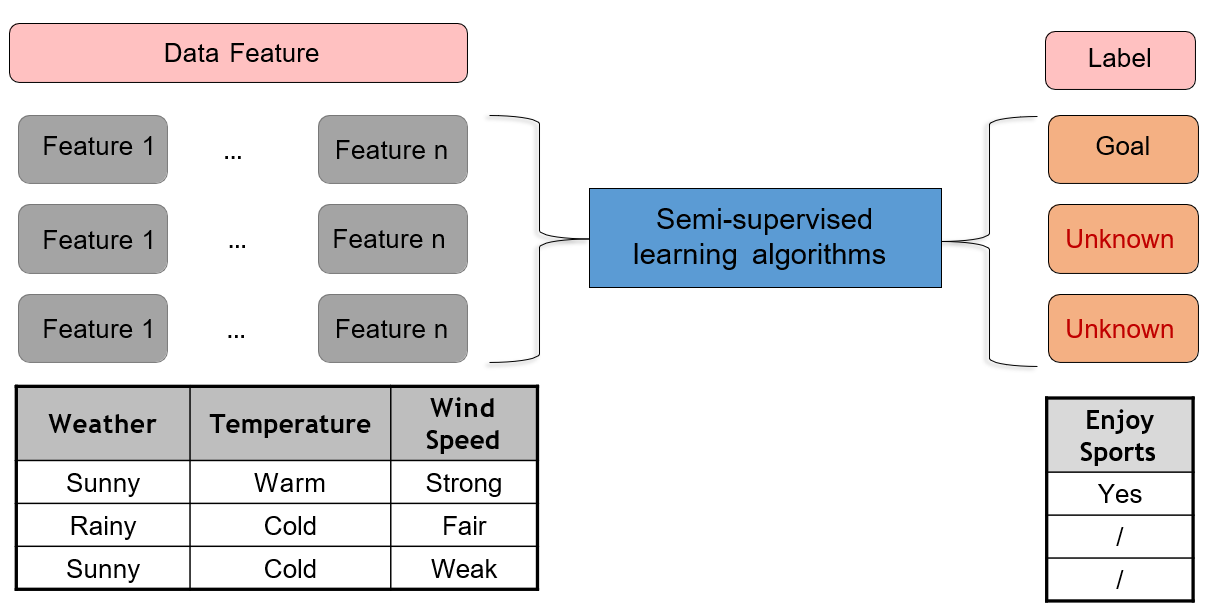
For unlabeled samples, the learning algorithms directly model the input datasets. Clustering is a common form of unsupervised learning. We only need to put highly similar samples together, calculate the similarity between new samples and existing ones, and classify them by similarity.



**Fig 3.5.2:** Unsupervised Machine Learning.

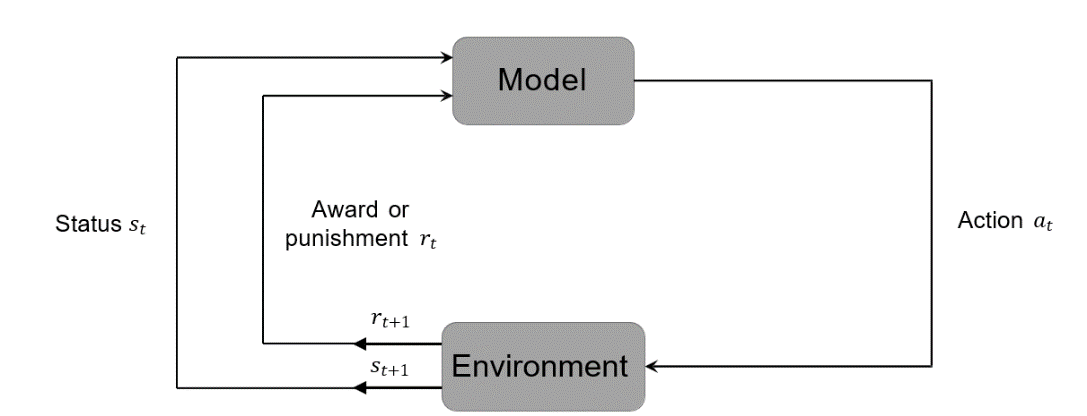
* *Semi-supervised learning:*

In one task, a machine learning model that automatically uses a large amount of unlabelled data to assist learning directly of a small amount of labelled data.



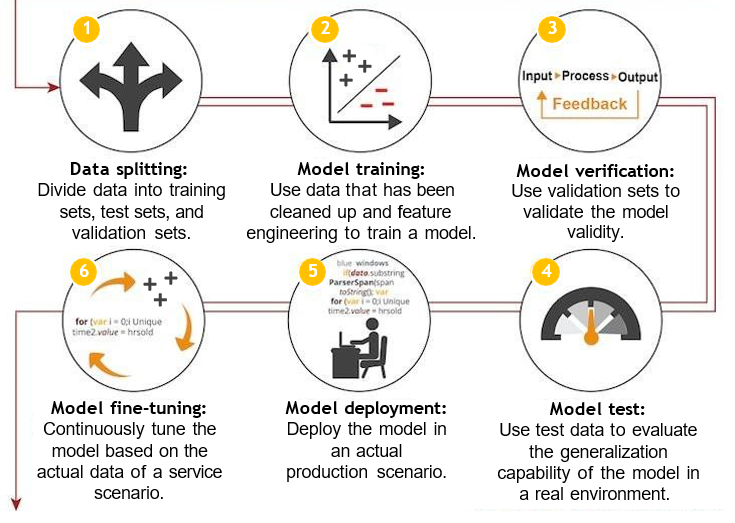
**Fig 3.5.4:** Semi-supervised Machine Learning.

* *Reinforcement learning:*

****It is an area of machine learning concerned with how agents ought to take actions in an environment to maximize some notion of cumulative reward. The difference between reinforcement learning and supervised learning is the teacher signal. The reinforcement signal provided by the environment in reinforcement learning is used to evaluate the action (scalar signal) rather than telling the learning system how to perform correct actions.

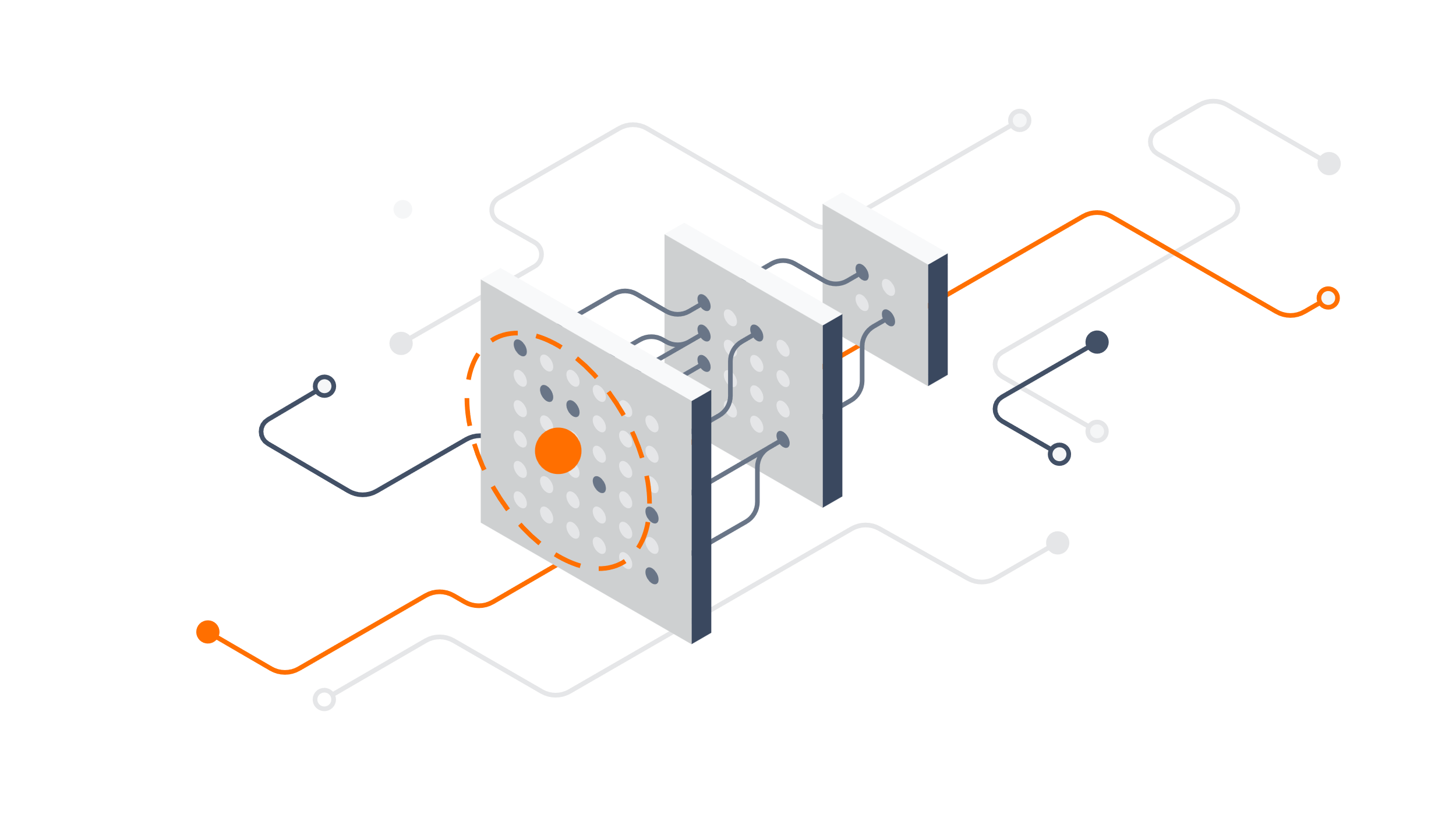
**Fig 3.5.5:** Reinforcement Learning.

**3.6 Procedure of Machine Learning**

The basic procedure of Model building through Machine Learning algorithm can be understood with the help of following flowchart:

**Fig 3.6.1:** Machine Learning Process

Each mentioned step has its own significance for machine learning process and can affect the accuracy & efficiency of model if not configured correctly.

**Chapter 4**

**Deep Learning**

*Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to “learn” from large amounts of data.*

**4.1 About Deep Learning**

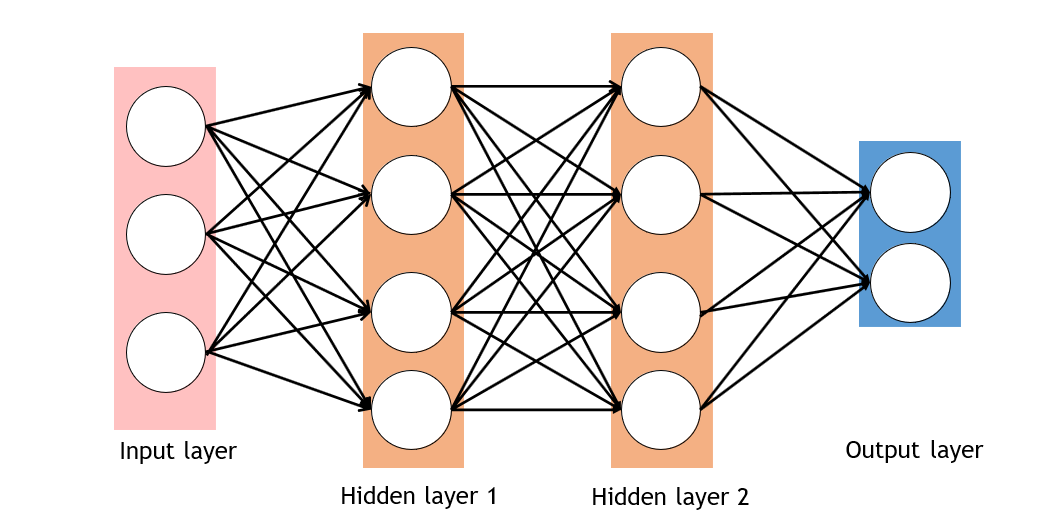
Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep learning is a key technology behind driverless cars, enabling them to recognize a stop sign, or to distinguish a pedestrian from a lamppost. It is the key to voice control in consumer devices like phones, tablets, TVs, and hands-free speakers. Deep learning is getting lots of attention lately and for good reason. It’s achieving results that were not possible before.

In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers.

**4.2 Neural Networks**

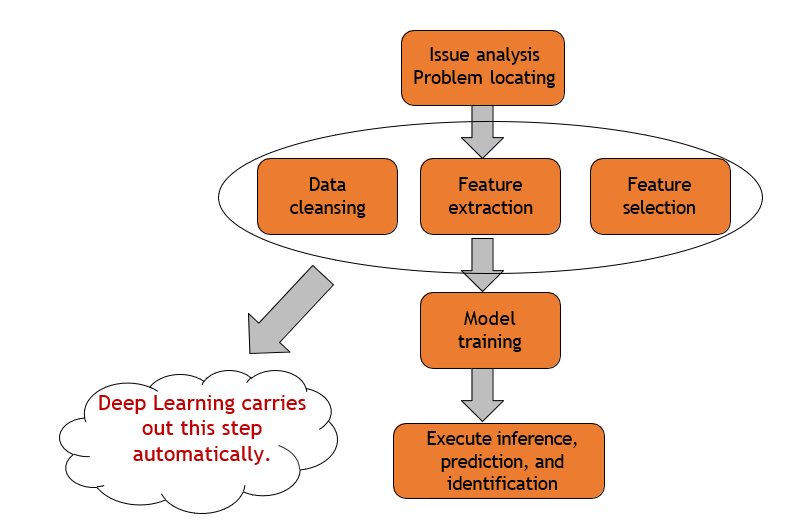
A neural network as a computer system composed of simple and highly interconnected processing elements, which process information by dynamic response to external inputs. A neural network can be simply expressed as an information processing system designed to imitate the human brain structure and functions based on its source, features, and explanations.

Artificial neural network (neural network): Formed by artificial neurons connected to each other, the neural network extracts and simplifies the human brain's microstructure and functions. It is an important approach to simulate human intelligence and reflect several basic features of human brain functions, such as concurrent information processing, learning, association, model classification, and memory.



**Fig 4.2.1:** Neural Network Representation.

**4.3 Difference – Machine Learning & Deep Learning**

As a model based on unsupervised feature learning and feature hierarchy learning, deep learning has great advantages in fields such as computer vision, speech recognition, and natural language processing. But some key points of difference can be mentioned with as:

**Fig 4.3.1:** Deep Learning advantage.

|  |  |
| --- | --- |
| **Traditional Machine Learning** | **Deep Learning** |
| Low hardware requirements on the computer: Given the limited computing amount, the computer does not need a GPU for parallel computing generally. | Higher hardware requirements on the computer: To execute matrix operations on massive data, the computer needs a GPU to perform parallel computing. |
| Applicable to training under a small data amount and whose performance cannot be improved continuously as the data amount increases. | The performance can be high when high- dimensional weight parameters and massive training data are provided. |
| Level-by-level problem breakdown | E2E learning |
| Manual feature selection | Algorithm-based automatic feature extraction |
| Easy-to-explain features | Hard-to-explain features |

As mentioned, we can observe that Deep Learning is more efficient as compared to Machine Learning in terms of model optimization.

**Chapter 5**

**Modules & Libraries**

**5.1 Python Modules**

A Python module is a file containing Python definitions and statements. A module can define functions, classes, and variables. A module can also include runnable code. Grouping related code into a module makes the code easier to understand and use. It also makes the code logically organized.

**5.2 Python Libraries**

A Python library is a collection of related modules. It contains bundles of code that can be used repeatedly in different programs. It makes Python Programming simpler and convenient for the programmer. As we don’t need to write the same code again and again for different programs. Python libraries play a very vital role in fields of Machine Learning, Data Science, Data Visualization, etc.

Some popular Python Libraries are as follows:

* **NumPy** is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.
* **SciPy** is a free and open-source Python library used for scientific computing and technical computing. SciPy contains modules for optimization, linear algebra, integration, interpolation, special functions, FFT, signal and image processing, ODE solvers and other tasks common in science and engineering.
* **Matplotlib** is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK.
* **Scikit-learn** is a machine learning library for the Python programming language.[3] It features various classification, regression and clustering algorithms including support-vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.
* **TensorFlow** is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks.
* **Keras** is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Up until version 2.3, Keras supported multiple backends, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML.
* **Librosa** is a Python package for music and audio analysis. Librosa is basically used when we work with audio data like in music generation (using LSTM's), Automatic Speech Recognition. It provides the building blocks necessary to create the music information retrieval systems.

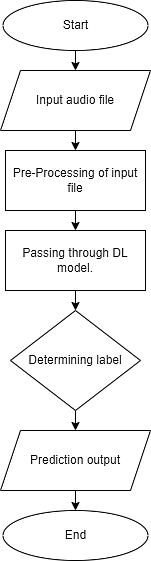
**Chapter 6**

**Project Description**

**6.1 Problem Statement**

The main purpose of the project is to return the emotion of a voice recording passed through it.

* 1. **Workflow of Project**



* Input audio file/recording.
* Normalization of input file.
* Passing of file through trained model.
* Label determination based on output of model.
* Printing of output.

**Fig 6.2.1:** Workflow

* 1. **Tech Stack Used**

For the development of the project following technologies has been used:

* Python – Deep Learning.
* TensorFlow.
* Keras.
* Librosa.
* Matplotlib.
* Scikit-Learn.
* Pandas.
* NumPy.
* Pickle.

The project is developed with the help of Jupyter Lab. Also the hardware dependencies are as:

* Windows 11 OS based Workstation.
* Intel i5 Processor.
* Nvidia GTX 1650 – Cuda & Cudnn Enabled.
* 10 GB Free Hard Disk space for Dataset.
* 4 GB Free Ram.
  1. **Dataset Used**
* *Toronto emotional speech set (TESS)*

There are a set of 200 target words were spoken in the carrier phrase "Say the word \_' by two actresses (aged 26 and 64 years) and recordings were made of the set portraying each of seven emotions (anger, disgust, fear, happiness, pleasant surprise, sadness, and neutral). There are 2800 data points (audio files) in total.

The dataset is organized such that each of the two female actor and their emotions are contain within its own folder. And within that, all 200 target words audio file can be found. The format of the audio file is a WAV format

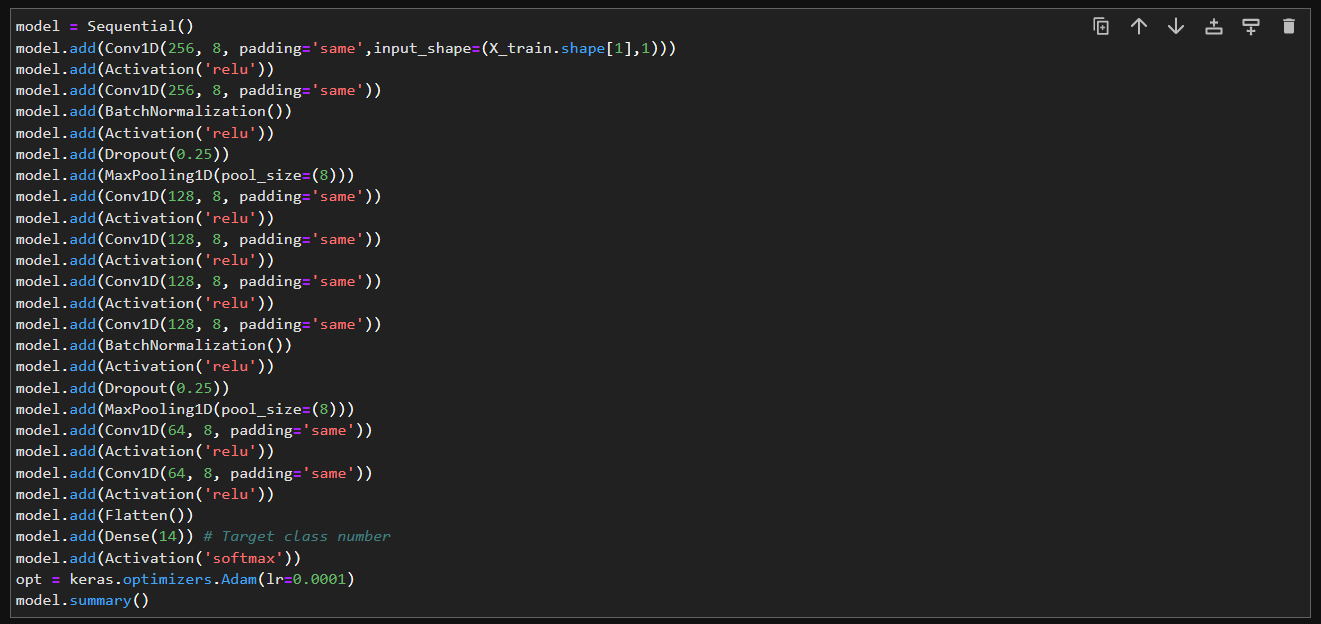
* *Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS)*

RAVDESS contains 1440 files: 60 trials per actor x 24 actors = 1440. The RAVDESS contains 24 professional actors (12 female, 12 male), vocalizing two lexically-matched statements in a neutral North American accent. Speech emotions includes calm, happy, sad, angry, fearful, surprise, and disgust expressions. Each expression is produced at two levels of emotional intensity (normal, strong), with an additional neutral expression.

* *Surrey Audio-Visual Expressed Emotion (SAVEE)*

Surrey Audio-Visual Expressed Emotion (SAVEE) database has been recorded as a pre-requisite for the development of an automatic emotion recognition system. The database consists of recordings from 4 male actors in 7 different emotions, 480 British English utterances in total. The sentences were chosen from the standard TIMIT corpus and phonetically-balanced for each emotion. The data were recorded in a visual media lab with high quality audio-visual equipment, processed and labeled. To check the quality of performance, the recordings were evaluated by 10 subjects under audio, visual and audio-visual conditions. Classification systems were built using standard features and classifiers for each of the audio, visual and audio-visual modalities, and speaker-independent recognition rates of 61%, 65% and 84% achieved respectively.

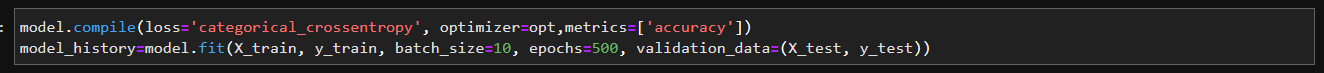
* 1. **Significance of important code segments.**
* *Model Definition.*

****

**Fig 6.5.1:** Model Definition

In this segment of code, the layer of Sequential model is defined made up with combination of Convo1D, MaxPooling, Dense, Normalization, Flatten, Activation.

* *Model Compilation.*

**

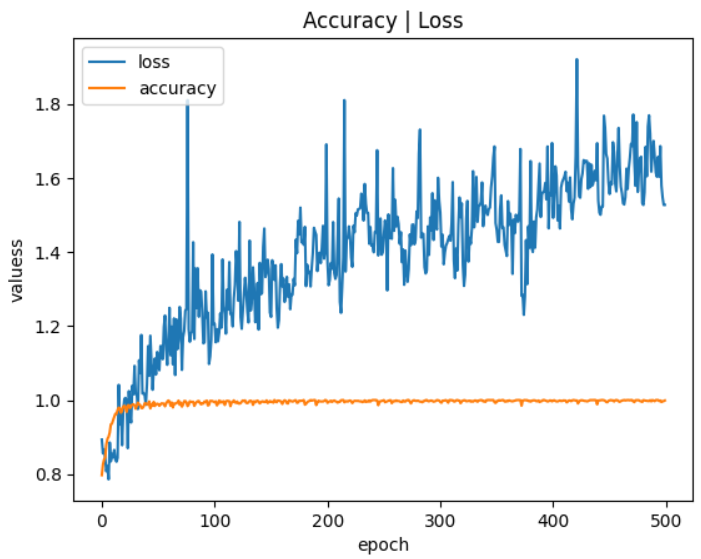
**Fig 6.5.2:** Model Compilation

In this segment of code model compilation is done with Adam Optimization, Categorical Entropy Loss over Accuracy Metrices. And 500 epochs are defined with a batch size of 10.

* 1. **Accuracy & Loss**

A loss function is used to optimize a machine learning algorithm. The loss is calculated on training and validation and its interpretation is based on how well the model is doing in these two sets. It is the sum of errors made for each example in training or validation sets. Loss value implies how poorly or well a model behaves after each iteration of optimization.

An accuracy metric is used to measure the algorithm’s performance in an interpretable way. The accuracy of a model is usually determined after the model parameters and is calculated in the form of a percentage. It is the measure of how accurate your model's prediction is compared to the true data.



**Fig 6.6.1:** Accuracy & Loss representation.

**References**

* Python 3.10.9 Documentation.
* Librosa Documentation.
* TensorFlow Documentation.
* Kaggle.
* Research Paper: Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology.