

**Industrial Internship Report on**  
**” Prediction of Agriculture Crop Production in India”**

**Prepared by**

**[Hema Varshini Loya]**

<i>Executive Summary</i>
<p>This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT).</p> <p>This internship was focused on a project/problem statement provided by UCT. We had to finish the project including the report in 6 weeks’ time.</p> <p>My project was Prediction of Agriculture Crop Production in India</p> <p>This internship gave me a very good opportunity to get exposure to Industrial problems and design/implement solution for that. It was an overall great experience to have this internship.</p>

**TABLE OF CONTENTS**

1	Preface .....	3
2	Introduction .....	5
2.1	About UniConverge Technologies Pvt Ltd .....	5
2.2	About upskill Campus .....	9
2.3	Objective .....	11
2.4	Reference .....	11
2.5	Glossary.....	11
3	Problem Statement.....	13
4	Existing and Proposed solution.....	14
5	Proposed Design/ Model .....	15
5.1	High Level Diagram (if applicable) .....	16
5.2	Low Level Diagram (if applicable) .....	16
5.3	Interfaces (if applicable) .....	17
6	Performance Test.....	19
6.1	Test Plan/ Test Cases .....	19
6.2	Test Procedure .....	20
6.3	Performance Outcome .....	20
7	My learnings.....	21
8	Future work scope .....	22

## 1 Preface

During the first week of the Data Science & Machine Learning internship program, our focus was on providing a comprehensive overview of the field, introducing the interns to the core concepts of data science and machine learning. The week was designed to lay a strong foundation and equip the interns with the necessary knowledge to tackle real-world challenges in the subsequent weeks.

During the second week of the Data Science & Machine Learning internship program, our focus was on providing a comprehensive overview of the field, introducing the interns to the core concepts of data science and machine learning and big data. The week was designed to lay a strong foundation on data science, impact on big data on business and data analyst vs data scientists.

During the third week of the Data Science & Machine Learning internship program, our focus was on providing a comprehensive overview of the field, introducing the interns to the core concepts data science and machine learning. The week was designed to lay a strong foundation on AI vs data science, introduction to probability and statistics and data science skills.

During the fourth week of the Data Science & Machine Learning internship program, our focus was on providing a comprehensive overview of the field, introducing the interns to the core concepts data science and machine learning. The week was designed to lay a strong foundation on machine learning, linear function in data science, optimization techniques in machine learning.

During the fifth week of the Data Science & Machine Learning internship program, our focus was on providing a comprehensive overview of the field, introducing the interns to learn about success ladder to the corporate world, how much a data engineer may earn and Data science interview question. The week was designed to lay a strong foundation on machine learning, linear function in data science, optimization techniques in machine learning.

During the sixth week of the Data Science & Machine Learning internship program, our focus was on providing a comprehensive overview of the field, introducing the interns to the core concepts data science and machine learning. The week was designed to work on project.

About need of relevant Internship in career development.

Internships can provide a great transition between the education process and a career. There is a tradeoff involved with internship positions, however. On the one hand, many internships are unpaid. On the other hand, you get an apprentice-level experience not typically available. By bridging the gap between learning and finding a career, internships can speed up the process of moving you toward your career goals.

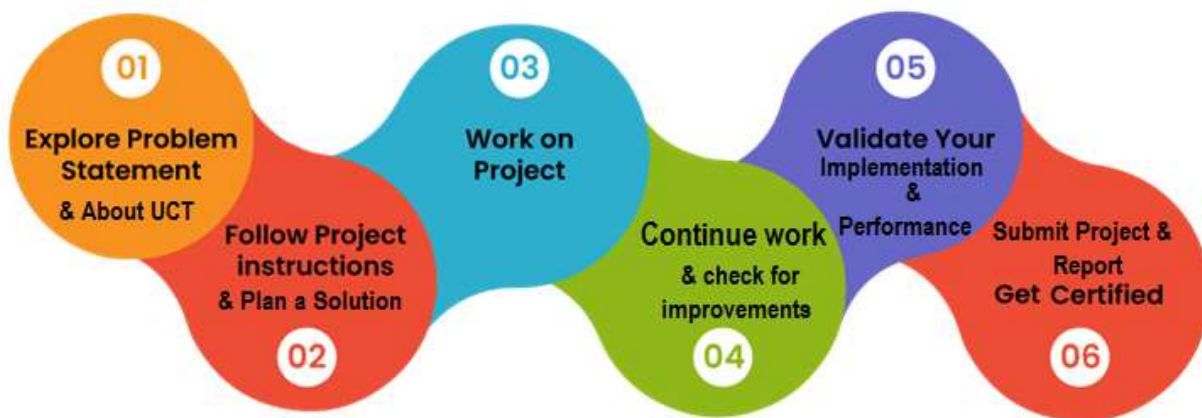
Brief about Your project/problem statement.

Across the Globe India Is the Second Largest Country having People more than 1.3 billion. Many People are dependent on the agriculture and it is the main resource. In agriculture cultivation/Production Having More Problems. I want to solve the big problem in India and useful to many more people. To predict Agriculture Production in India from 2004-2012.

Opportunity given by USC/UCT.

USC offers exciting and diverse opportunities to spend a semester or year studying abroad through programs offered in over 50 cities on five continents. Some programs include opportunities for research, experiential learning or an internship. All students are encouraged to integrate study abroad into their overall USC experience.

How Program was planned



Your Learnings and overall experience.

My learning and experience with upskill campus was magnificent. I have learn lot of thing for the internship.

Thank to all (the admin), who have helped you directly or indirectly.

Your message to your juniors and peers.

upskill campus was a grate platform to do internship for any domain. My message to junior was only one that you have to join this internship.

## 2 Introduction

### 2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies** e.g. **Internet of Things (IoT), Cyber Security, Cloud computing (AWS, Azure), Machine Learning, Communication Technologies (4G/5G/LoRaWAN), Java Full Stack, Python, Front end** etc.



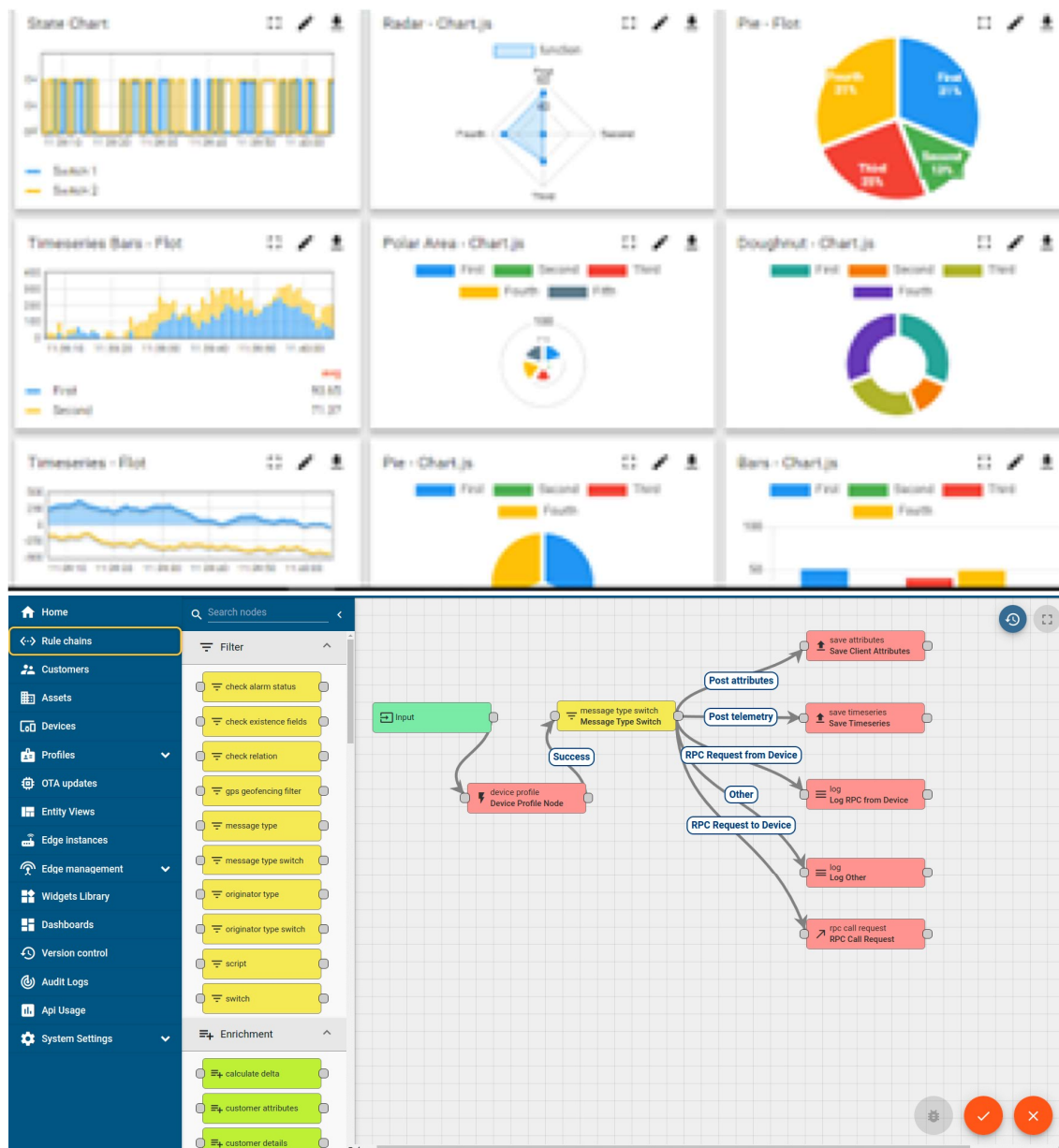
#### i. UCT IoT Platform ( )

**UCT Insight** is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

It has features to

- Build Your own dashboard
- Analytics and Reporting
- Alert and Notification
- Integration with third party application(Power BI, SAP, ERP)
- Rule Engine





## ii. Smart Factory Platform ( **FACTORY WATCH** )

Factory watch is a platform for smart factory needs.

It provides Users/ Factory

- with a scalable solution for their Production and asset monitoring
- OEE and predictive maintenance solution scaling up to digital twin for your assets.
- to unleash the true potential of the data that their machines are generating and helps to identify the KPIs and also improve them.
- A modular architecture that allows users to choose the service that they want to start and then can scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.



Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i





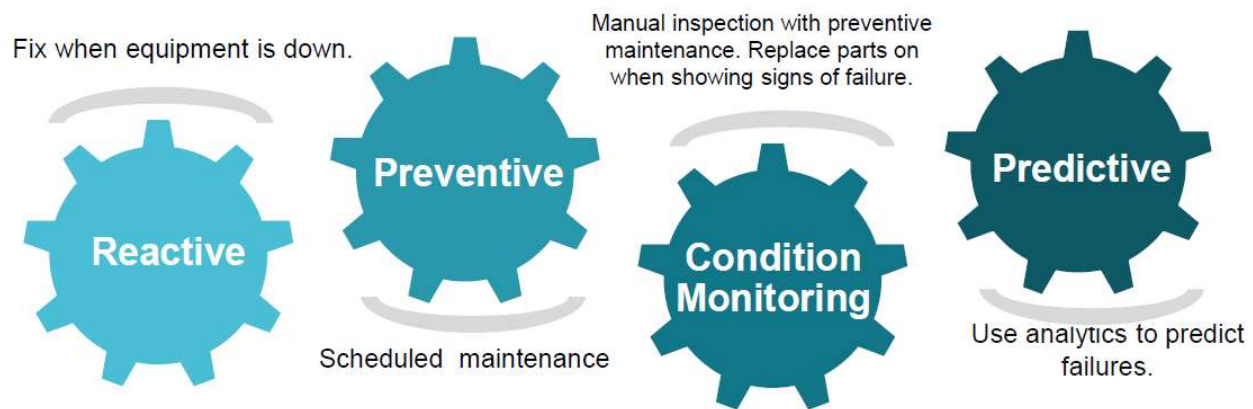


### iii. based Solution

UCT is one of the early adopters of LoRAWAN technology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

### iv. Predictive Maintenance

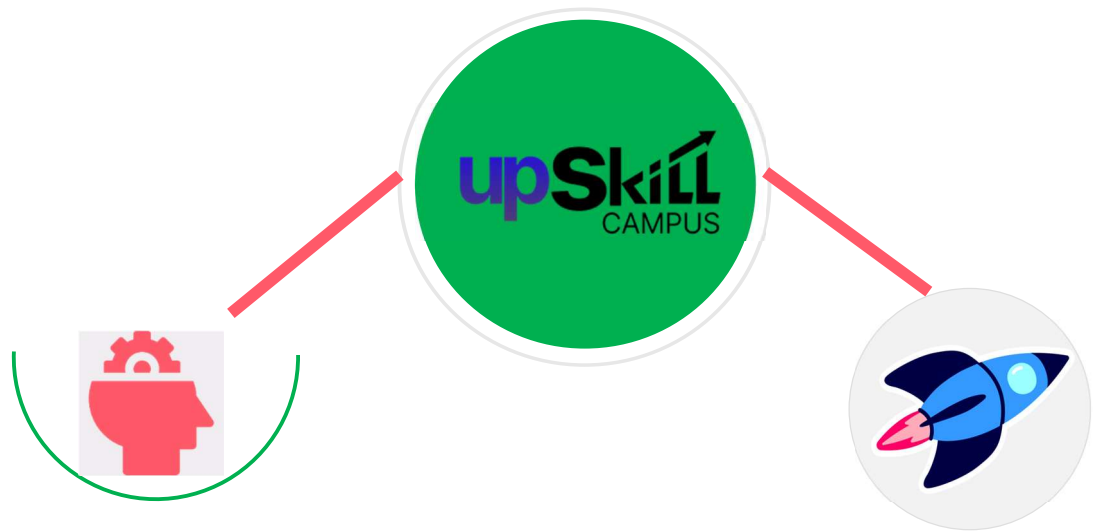
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



## 2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

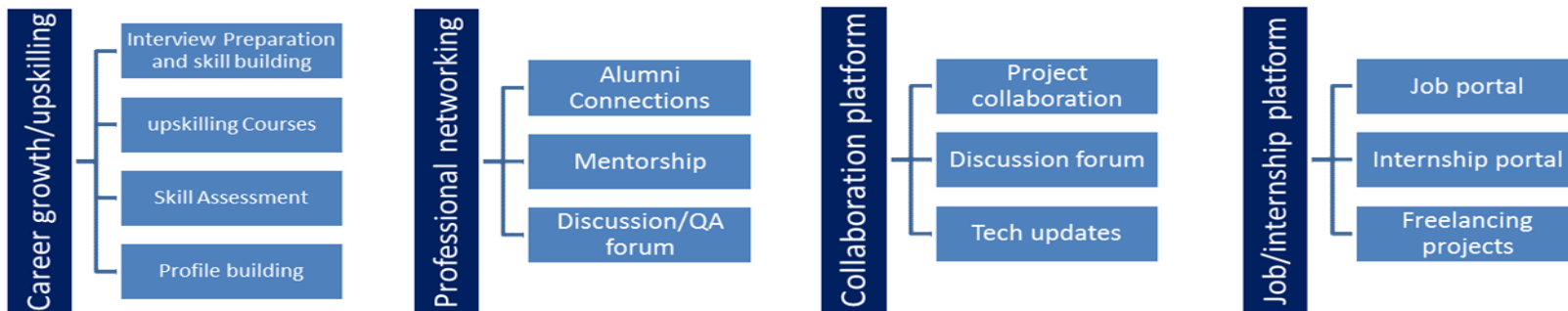
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services

upSkill Campus aiming to upskill 1 million learners in next 5 year

<https://www.upskillcampus.com/>



## 2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs in collaboration with EICT Academy, IITK, IITR and IITG in multiple domains.

## 2.4 Objectives of this Internship program

The objective for this internship program was to

- get practical experience of working in the industry.
- to solve real world problems.
- to have improved job prospects.
- to have Improved understanding of our field and its applications.
- to have Personal growth like better communication and problem solving.

## 2.5 Reference

- [1] <https://www.sciencedirect.com/science/article/pii/S0168169920302301#bb0005>
- [2] <https://www.mdpi.com/2076-3417/11/21/9816>
- [3] <https://www.hindawi.com/journals/complexity/2021/5534379/>

## 2.6 Glossary

Terms	Acronym
<b>accuracy</b>	The number of correct classification prediction divided by the total number of predictor
Confusion matrix	An NaN table that summarize the number of correct and incorrect prediction that a classification model made
Regression data	Regression feature are the continuous data
Linear Regression	A supervised model which predict the continuous data
matplotlib	An open source python 2D plotting library helps you visualize different aspect of machine learning

## 2.7 Code submission (Github link)

<https://github.com/hemavarshiniloya/Prediction-of-Agriculture-Crop-Production-in-India>

## 2.8 Report submission (Github link) : first make placeholder, copy the link.

<https://github.com/hemavarshiniloya/Prediction-of-Agriculture-Crop-Production-in-India>

### 3. Problem Statement

Problem Statement of my project is to predict the agriculture crop in the India.

Agriculture in India depends heavily on weather and climate conditions. Weather forecasts are useful for decisions regarding crop choice, crop variety, planting/harvesting dates, and investments in farm inputs such as irrigation, fertilizer, pesticide, herbicide etc. Hence, improved weather forecast based agroment advisory service greatly helps farmers to take advantage of benevolent weather and mitigate the impacts of malevolent weather situation.

Across the Globe India Is the Second Largest Country having People more than 1.3 billion. Many People are dependent on the agriculture and it is the main resource. In agriculture cultivation/Production Having More Problems. I want to solve the big problem in India and useful to many more people.

I am using machine learning algorithm to predict the crop production with the dataset of crop and years.

## 4.Existing and Proposed solution

### Existing solution

Machine learning is an important decision support tool for crop yield prediction, including supporting decisions on what crops to grow and what to do during the growing season of the crops. Several machine learning algorithms have been applied to support crop yield prediction research. In this study, we performed a Systematic Literature Review (SLR) to extract and synthesize the algorithms and features that have been used in crop yield prediction studies. Based on our search criteria, we retrieved 567 relevant studies from six electronic databases, of which we have selected 50 studies for further analysis using inclusion and exclusion criteria. We investigated these selected studies carefully, analyzed the methods and features used, and provided suggestions for further research. According to our analysis, the most used features are temperature, rainfall, and soil type, and the most applied algorithm is Artificial Neural Networks in these models. After this observation based on the analysis of machine learning-based 50 papers, we performed an additional search in electronic databases to identify deep learning-based studies, reached 30 deep learning-based papers, and extracted the applied deep learning algorithms. According to this additional analysis, Convolutional Neural Networks (CNN) is the most widely used deep learning algorithm in these studies, and the other widely used deep learning algorithms are Long-Short Term Memory (LSTM) and Deep Neural Networks (DNN).

### Proposed Solution

The proposed solution for my project i.e prediction of agriculture crop production in India. I am using the machine learning in the project to predict the crop. Using machine learning models like data collection, data cleaning, data visualization, model training. In the project, my am using Linear Regression algorithm to get the accuracy for the project.

### What value addition are you planning?

I am adding the data visualization, data cleaning and Linear Regression algorithm.



## 5. Proposed Design/ Model

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

*#Data Collection*

```
data=pd.read_csv("C:\\Users\\hemav\\Downloads\\datafile - datafile.csv")
```

In [3]: *#Data Preprocessing*  
data

Out[3]:

	Crop	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12
0	Rice	100	101	99	105	112	121	117	110
1	Wheat	100	101	112	115	117	127	120	108
2	Coarse Cereals	100	107	110	115	113	123	122	136
3	Pulses	100	108	134	124	124	146	137	129
4	Vegetables	100	109	103	118	113	124	128	115
5	Fruits	100	99	99	98	102	104	114	119
6	Milk	100	97	98	98	98	112	123	124
7	Eggs, Fish and Meat	100	102	101	100	99	116	133	137
8	Oilseeds	100	86	85	97	104	103	99	102
9	Sugarcane	100	96	91	87	80	81	109	107
10	Fibers	100	92	91	96	109	107	138	140
11	All Agriculture	100	99	101	104	106	115	123	122

In [8]: *#checking the null values*  
data.isnull().sum()

Out[8]: Crop 0  
2004-05 0  
2005-06 0  
2006-07 0  
2007-08 0  
2008-09 0  
2009-10 0  
2010-11 0  
2011-12 0  
dtype: int64

## 5 High Level Diagram (if applicable)

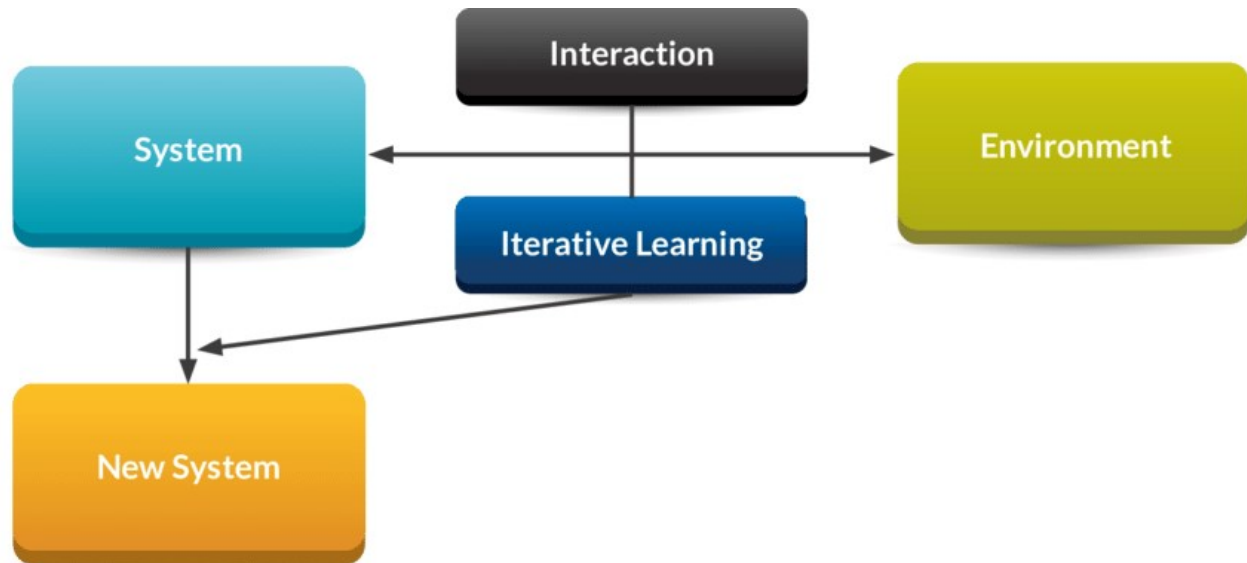
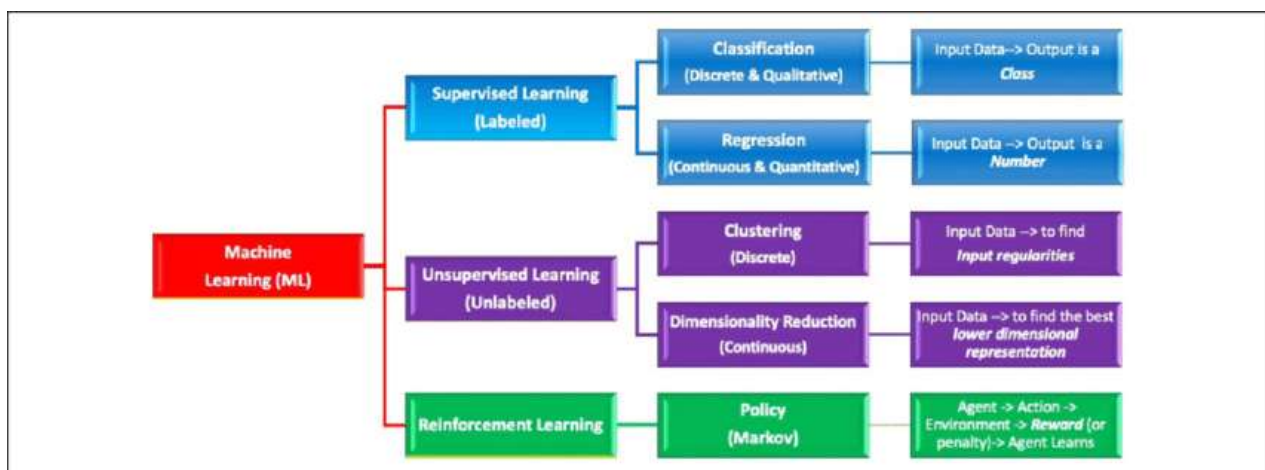


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

### 5.1 Low Level Diagram (if applicable)

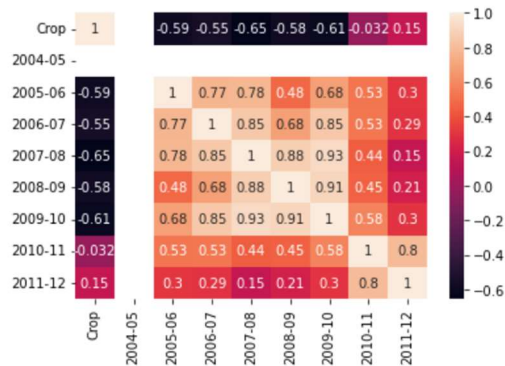


## 5.2 Interfaces (if applicable)

Update with Block Diagrams, Data flow, protocols, FLOW Charts, State Machines, Memory Buffer Management.

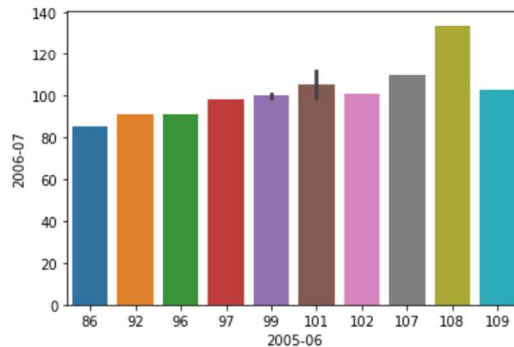
```
In [14]: sns.heatmap(data.corr(),annot=True)
```

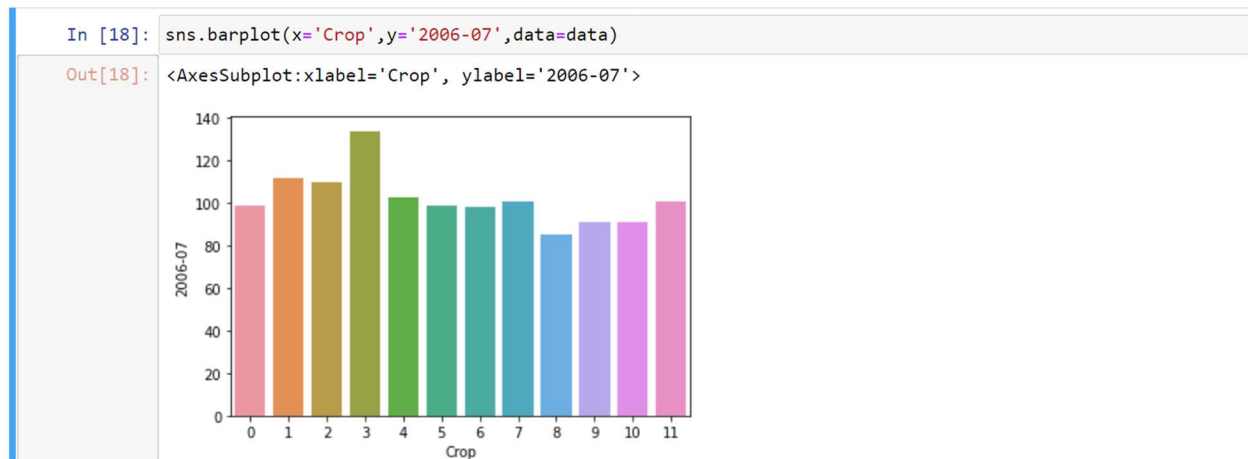
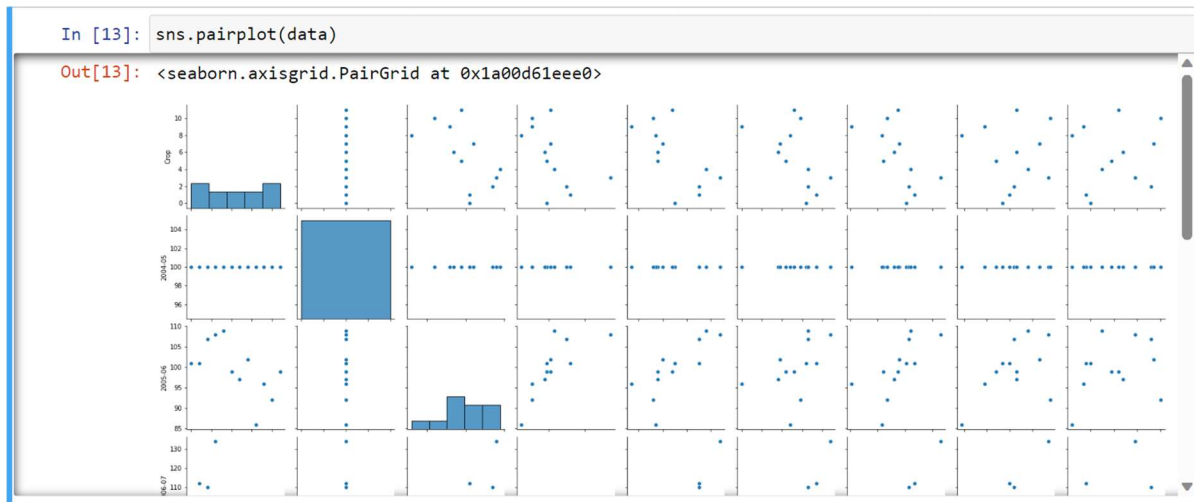
```
Out[14]: <AxesSubplot:>
```



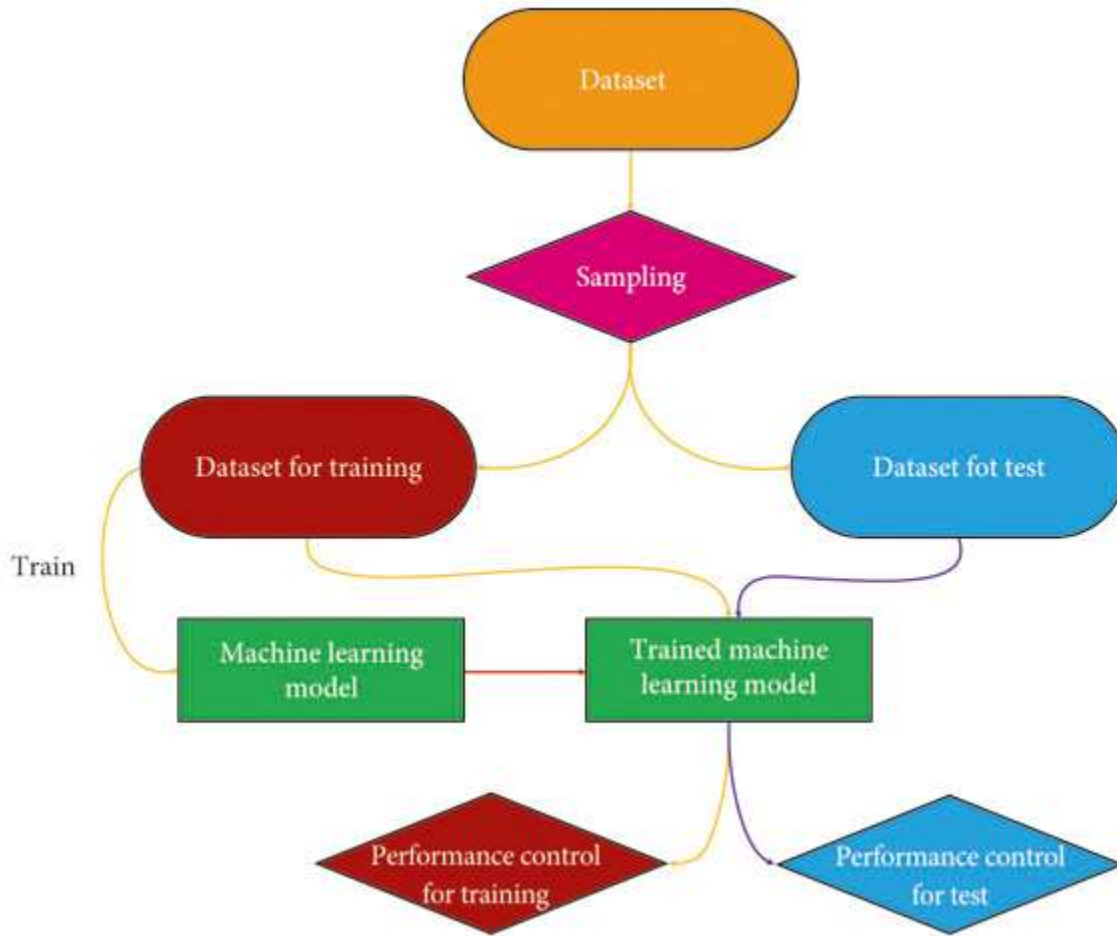
```
In [23]: sns.barplot(x='2005-06',y='2006-07',data=data)
```

```
Out[23]: <AxesSubplot:xlabel='2005-06', ylabel='2006-07'>
```





## 6 Performance Test



### 6.1 Test Plan/ Test Cases

```
In [48]: #importing the split function
from sklearn.model_selection import train_test_split
```

```
In [49]: #Loading the feature values and target value in the split function
lr_train,lr_test,llr_train,llr_test=train_test_split(lr,llr,test_size=0.2,shuffle=True)
```

## 6.2 Test Procedure

```
In [50]: #importing the algorithm  
from sklearn.linear_model import LinearRegression
```

```
In [51]: model=LinearRegression()
```

```
In [53]: model.fit(lr,llr)
```

```
Out[53]: LinearRegression()
```

```
In [54]: prediction=model.predict(lr_test)
```

## 6.3 Performance Outcome

```
In [55]: from sklearn.metrics import r2_score
```

```
In [56]: r2_score(llr_test,prediction)
```

```
Out[56]: 0.8686426145779835
```



## 7 My learnings

You should provide summary of your overall learning and how it would help you in your career growth.

Form the internship I have learnt about the data science and machine learning which is a best career for the upcoming days. By using data science and machine learning I gain understanding about problem solving and analysis. It is used in my career growth by solving the big problems and understanding the problem etc.

## 8 Future work scope

Machine learning is a rapidly evolving field with a vast range of applications. As we look into the future, there are several exciting areas where machine learning is expected to make significant contributions. Here are some potential work scopes in machine learning:

1. **Deep Learning Advances:** Deep learning has already achieved remarkable success in various domains, but there is still room for improvement. Future work will focus on developing more powerful neural network architectures, improving training algorithms, and addressing challenges such as interpretability and robustness.
2. **Reinforcement Learning:** Reinforcement learning (RL) enables machines to learn optimal behaviors through trial and error. Future work in RL will likely focus on developing more efficient algorithms that can handle complex tasks with sparse rewards, improving sample efficiency, and exploring real-world applications in areas like robotics and autonomous systems.
3. **Explainable AI:** Interpreting and understanding the decisions made by machine learning models is crucial, especially in sensitive domains such as healthcare and finance. Future work will involve developing techniques to make AI models more transparent, explainable, and accountable, enabling users to trust and understand the decisions made by these models.
4. **Federated Learning:** As data privacy concerns grow, federated learning has emerged as a promising approach that allows training machine learning models across multiple decentralized devices or organizations without sharing raw data. Future work will focus on developing more efficient and secure federated learning algorithms, addressing issues like data heterogeneity and model aggregation.
5. **Transfer Learning and Few-Shot Learning:** Transfer learning allows models to leverage knowledge learned from one task to improve performance on another related task. Future work will focus on developing more effective transfer learning methods, including techniques for domain adaptation, multi-task learning, and few-shot learning, where models can learn from only a few examples.
6. **Ethical and Fair AI:** Ensuring that machine learning models are fair, unbiased, and free from discrimination is a crucial area of research. Future work will involve developing techniques to detect and mitigate biases in datasets, creating fairness-aware algorithms, and addressing the ethical considerations surrounding the use of AI in various applications.
7. **AI for Healthcare:** Machine learning has the potential to revolutionize healthcare by enabling personalized medicine, disease diagnosis, drug discovery, and more. Future work will involve developing AI models that can integrate and analyze complex biomedical data, such as genomics, medical images, electronic health records, and wearable sensor data.

8. AI for Climate Change and Sustainability: Machine learning can play a significant role in addressing climate change and sustainability challenges. Future work will focus on developing models to optimize energy consumption, predict extreme weather events, analyze satellite imagery for environmental monitoring, and support decision-making for sustainable development.

9. AI and Natural Language Processing: Natural Language Processing (NLP) has seen significant advancements, but there is still room for improvement in areas like language understanding, generation, and machine translation. Future work will involve developing models that can comprehend and generate human-like text, understand context and semantics, and facilitate more natural human-computer interactions.

10. Quantum Machine Learning: With the development of quantum computing, there is growing interest in exploring the intersection of quantum computing and machine learning. Future work will focus on developing quantum machine learning algorithms and models that can harness the power of quantum computing to solve complex problems more efficiently.