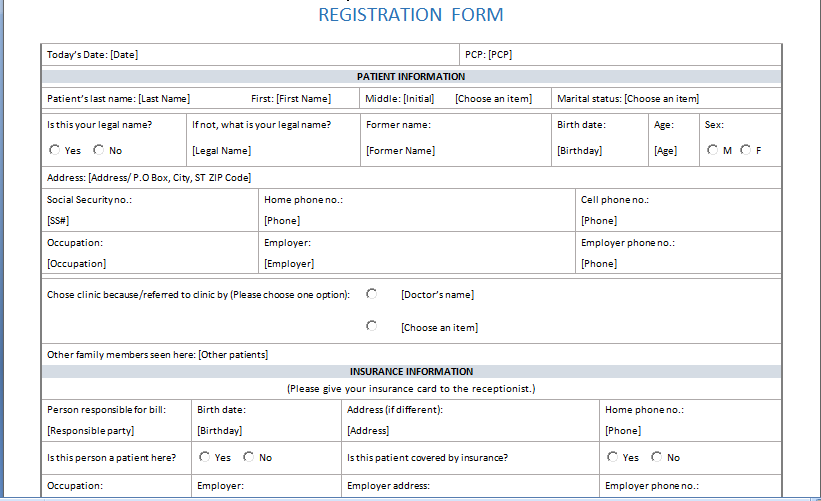
**Hospitalization Project**

**Q1. Tell me about your current project.**

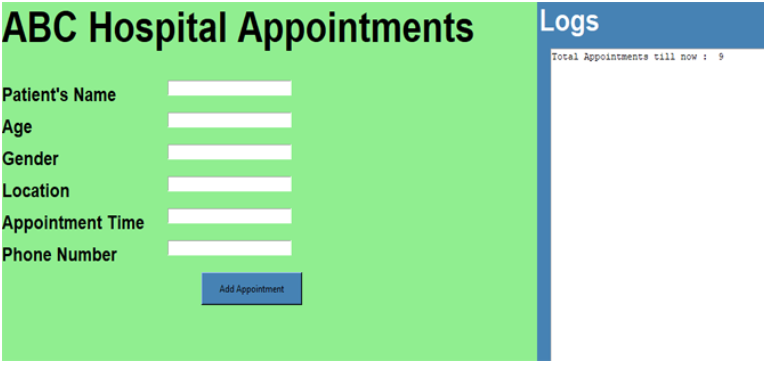
**Ans:** The Project name is Hospitalization.

The project goal of Hospitalization is to register outer/inner patients accordingly and guidance with the proper Doctor and Medical session and Tests.



Features of Hospital.

* Provides bed-related inpatient functions.
* Outpatient – related functions.
* Diagnostic and treatment functions.
* Administrative functions.
* Service function(food, supply)
* Research and teaching function.



Functionalities in the Hospitalization System

* Registration of Patients in the system.
* Registration required patient Name, Gender, Age, DOB, Marital status.
* Patient referral or appointment scheduling(time in , timeout)
* Preparing Bill will help to continue the process

**Deployment/Installation**

* The application can be easily installed as a web-based API on any cloud platform. This installation is similar to a plug and play scenario

**Q2. What was the size of the data?**

**Answer:** It is csv data of 1lakh rows

**Q3. What was the data type?**

**Answer:** The data used for training this model consisted of thousands of images; the images then are converted to tensor objects, which have a float 32 representation.

**Q4. What was the team size and distribution?**

**Answer:** Team consist of

* 1 Manager
* 1 Design Architect
* 1 Lead
* 2 Dev – Ops Engineer
* 2 UI developer
* 2 Data Scientist

**Q5.What Hadoop distribution were you using?**

**Answer:**

Hadoop splits the data and gives splits of data to several Map parts (Mappers). The resulting action rules from all the Mappers are combined in such a way that the action rule acts as a key and the support and confidence from all the Mappers acts as iterator list of values

The combined action rules are given to the Reduce part, where we propose using a Random Forest type of algorithm in order to combine the output from all the Mappers. The Random Forest algorithm works in analogy to checking, where if more than 50% of the patients is accepted. In our proposed implementation, the Random Forest algorithm checks the output from all the Mappers

**Q6.What is the version of distribution?**

**Answer:** CDH – 5.8.0

**Q7.What was the size of the cluster?**

**Answer:**

The cluster (production setup) consisted of 10 servers with

* Intel i7 processors
* 4 GB of RAM
* 150 GB of Secondary storage each
* Mounted NAS locations

**Q8. How many nodes were there in all the Dev, UAT, and Prod environments?**

**Answer:** The main application coding part will be in individual server also to come with speed respond we use other servers as well connecting and responding to cloud also need to maintain storage space as well.

In total, we had:

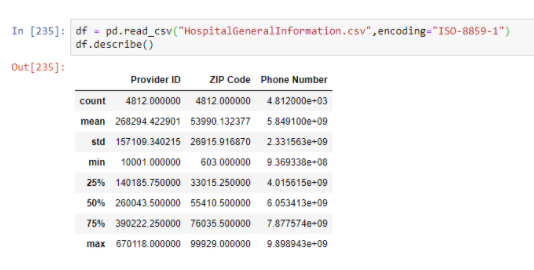
* 2 nodes in the dev environment,
* 2 nodes in UAT, and
* 5 nodes in production.

**Q9.How were you creating and maintaining the logs?**

**Answer:** Logins are taken care by Mongodb once application opened registration of the patient’s time in till the time out, personal details, inside/outside patient details. Medication and Test details are also noted down by Hospital through patient ID. Each model has been designed as individual method with Try and Exception block

**Q10.What techniques were you using for data pre-processing for various data science use cases and visualization?**

**Answer:** When we do the data preprocessing, cleaning, validation. We will find some NA missing values, blank, duplicates data which we will remove in the below steps.

****

* Preparing data model, data undergo multiple steps. So removing duplicates, missing values, blank will be part of our data

# Import pandas library

Import pandas as pd

#list of Tuples

Employees = [(‘Stuti’, 28, ‘varanasi’),

(‘saumya’, 32, ‘Delhi’),

(‘Aaditya’, 25,’Mumbai’),

(‘Saumya’, 32,’Delhi’),

(‘Saumya’,32,’Delhi’),

(‘Saumya’,32,’Mumbai’),

(‘Aaditya’, 40, ‘Dehradun’),

]

#creating a Database object

Df = pd.DataFrame(employees,

Columns = [‘Name’, ‘Age’,’City’])

#selecting duplicate rows except first

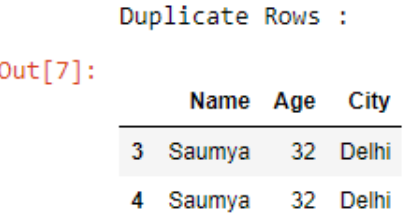
#occurence based on all columns

duplicate = df[df.duplicated([‘Name’,’Age])]

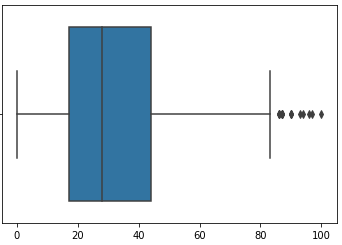
Print(“Duplicate Rows based on Name and age :”)

#print the resultant dataframe

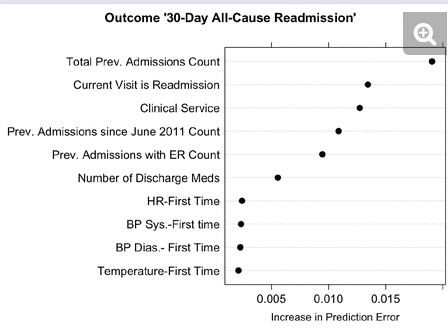
duplicate

****

* **Removing outliers**

****

**In the below fig clear shows how the input and output of patients increase in count for 30 days.**

****

**For Deep Learning**

A random forest is a machine learning technique that's used to solve regression and classification problems

**Q11. How were you maintaining the failure cases?**

**Answer:**

The specific problem(s) or issue(s) faced by stakeholders like patients that would be addressed by means of provision of improved services through the proposed project. When we write a code and start checking on loading data and o

n outcomes. If anyone reenter same patient name it can’t take. It will display already registered.

**Q12.What kind of automation has you done for data processing?**

**Answer:** The registration data entered in the Hospital system with patient Id along with Doctor ID once consultation done and prescription given by doctor automatically it will be updated in Medicine department.

**Q13.Have you used any scheduler?**

**Answer:**

Yes once in month retraining the model

**Q14.How are you monitoring your job?**

**Answer:** Creating a project and check on that by keeping on observation for 3 month and daily checks on when wrong data entity entered in the application Error occurs and guidance need able.

**Q15. What were your roles and responsibilities in the project?**

**Answer:** As we know best role will be given in the project but responsibilities give the task refer to Develops the Project Plan with the team and manages the team's performance of project tasks. Secures acceptance and approval of deliverables from the Project Sponsor and Stakeholders. Responsible for communication, including status reporting, risk management, escalation of issues that cannot be resolved in the team.

**Q16.What was your day to day task?**

**Answer:** My daily routine a collection of projects, with different priorities and expectations So my day-to-day work is a collection of projects, with different priorities and expectations. Some are short term, a day or two to complete; others have no set end, but will likely take a year or more as they are lower priority. A project is a specific group of tasks which lead to a defined outcome

**Q17.In which area you have contributed the most?**

**Answer:** Most contributed areas are model saving, deployment, data batches for prediction, data validation and Data transformation part. More than one project I involved myself.

**Q18.In which technology you are most comfortable?**

**Answer:**

Artificial intelligence (AI) is the technology used for equipping computer systems with the ability to make decisions like humans. Machine Learning, Deep Learning and Natural Language Processing really prefer to work on.

**Q19.How you rate yourself in big data technology?**

**Answer:** Big data technologies I can say one of the part of project I worked on opportunity get means have a knowledge to handle that.

**Q20. In how many projects you have already worked?**

**Answer**: I have worked on many small and large project decision trees, sentimental analysis, many more project s includes NLP and DL ML.

**Q21. How were you doing deployment?**

**Answer:** The client want to deploy on cloud based and the data entry is going to be stored on the cloud application. We deploy project through an API

**Q22.What kinds of challenges have you faced during the project?**

**Answer:** Conflict and tension, Lack of trust. Trust is crucial to teamwork, and it starts with people knowing each other. Low engagement. Lack of transparency. No long-term thinking needs to take decision as quick as possible.

**Q23.What will be your expectations?**

**Answer:** Proactively managing expectations leads to successful projects, satisfied stakeholders, and engaged teams. This involves addressing expectations during project planning and execution, frequent communication, and a focus on realistic goals.

**Q24. What is your future objective?**

**Answer:** Proper Mindset and Balance, Commitment to Improved Physical Health.Career Passion and Personal Satisfaction. Financial Stability.

**Q25. Why are you leaving your current organization?**

**Answer:** My reason for job change is upward mobility and financial growth. I have worked with my current employer for the last 4 years and have grown well. I am grateful for my experience as I have imbibed a spectacular work ethic and in-depth industry knowledge. And need challenging work part.

**Q26. How did you do Data validation?**

**Answer:** Data validation is accidental or intentional mistake in a program can prevent it from functioning correctly or even corrupt the data. Most of the patients number can become delicacy and show error in the mid of the process (when doctor consultation done and move further for medicine department that time error issue may occur)through date and time line we prepare one code for the continuously use in the Hospital.

**Q27. How did you do Data enrichment?**

The view performing a look up at the time of collection and appending the contextual information into the log we will check in data preprocessing, in the cleaning step only so that in the final deploy should not come any issues.

**Q28. How would you rate yourself in machine learning?**

**Answer:** 7

**Q29. How would you rate your self in distributed computation?**

**Answer:** 7

**Q30. What are the areas of machine learning algorithms that you already have explored?**

Answer: I have explored many machine learning algorithms like Linear Regression, Logistic Regression, Decision tree, Random Forest.