1. INTRODUCTION

1.1 PROJECT OVERVIEW

Through this project we are creating a basic structure of social media database which could easily be connected to frontend interface.

Here we are managing the data of multiple users, their followers, interests and public activity on the social media platform which includes post likes ,comments , comment likes , hashtag followed ,bookmarks and many more...

Through our SQL query we have shown a clear cut description of connection and inter relation between different activities on social media.

Here we have tried to fetch and store the data in its true storage form into this database like (img /videos in url)

With our true efforts and research we have tried to give a simple and more exact view of a basic social interconnecting site.

1.2 PURPOSE

With social media, you can share information and ideas in a variety of ways. Different outlets allow you to publish your own ideas in writing, with pictures, or through videos and voice recordings; and you can also hyperlink your audience to interesting articles, pictures, and videos. The information you share can be either private or public. For example, you can email a private message to one person and broadcast a video to a global audience via YouTube. Always consider your communication situation (see page 68) before sharing information .we see some of the purposes of social media.
Social media can also serve as personal learning tools (see page 283). You can get updates about your friends and family, or learn about what's happening in your community and around the world. The immediacy of the information flow allows you to get the gist of the latest news almost as it happens. Today, breaking news is often broadcast via social media before traditional media like TV and newspapers are able to cover it in detail. Social media can also reveal public sentiment about the big issues of the day or, on a less urgent level, opinions about a new restaurant or movie.

2. IDEATION AND PROPOSED SOLUTION

2.1 Problem Statement Definition

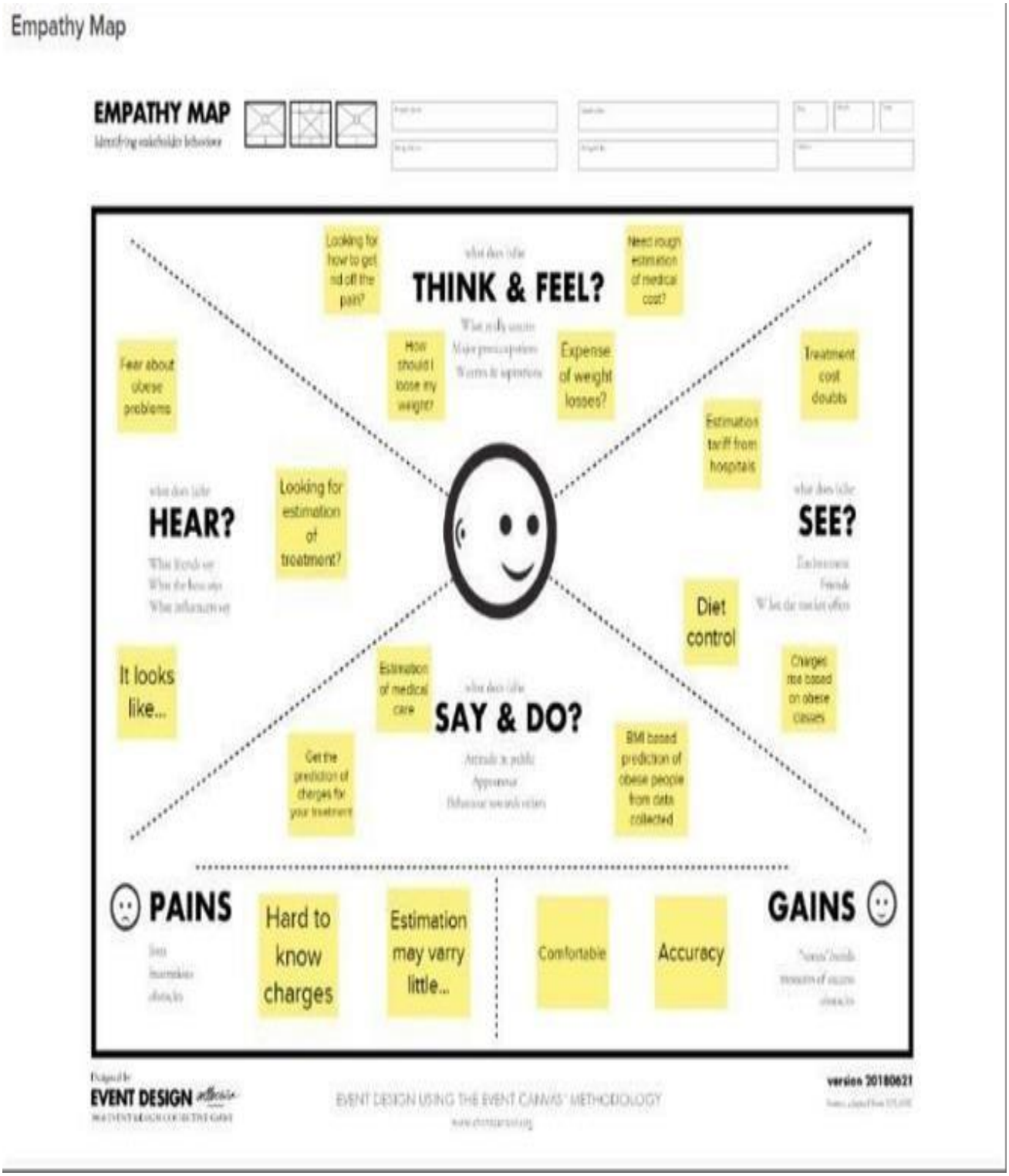
Create a problem statement to understand your customer's point of view. The Customer Problem Statement template helps you focus on what matters to create experiences people will love.

A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

2.2 Empathy Map Canvas

The empathy map canvas is a tool that helps project teams gain a deeper understanding of the users or stakeholders involved in the project. In the context of the breast cancer detection project, the empathy map canvas can be used to develop empathy towards the healthcare professionals who will be using the system and the patients who will benefit from early detection.

By completing the empathy map canvas, the project team gains valuable insights into the perspectives, needs, and expectations of the healthcare professionals and patients involved in breast cancer detection. This understanding helps inform the design, development, and implementation of the system, ensuring that it meets the users' requirements, aligns with their workflow, and delivers a positive impact on breast cancer detection and patient care.

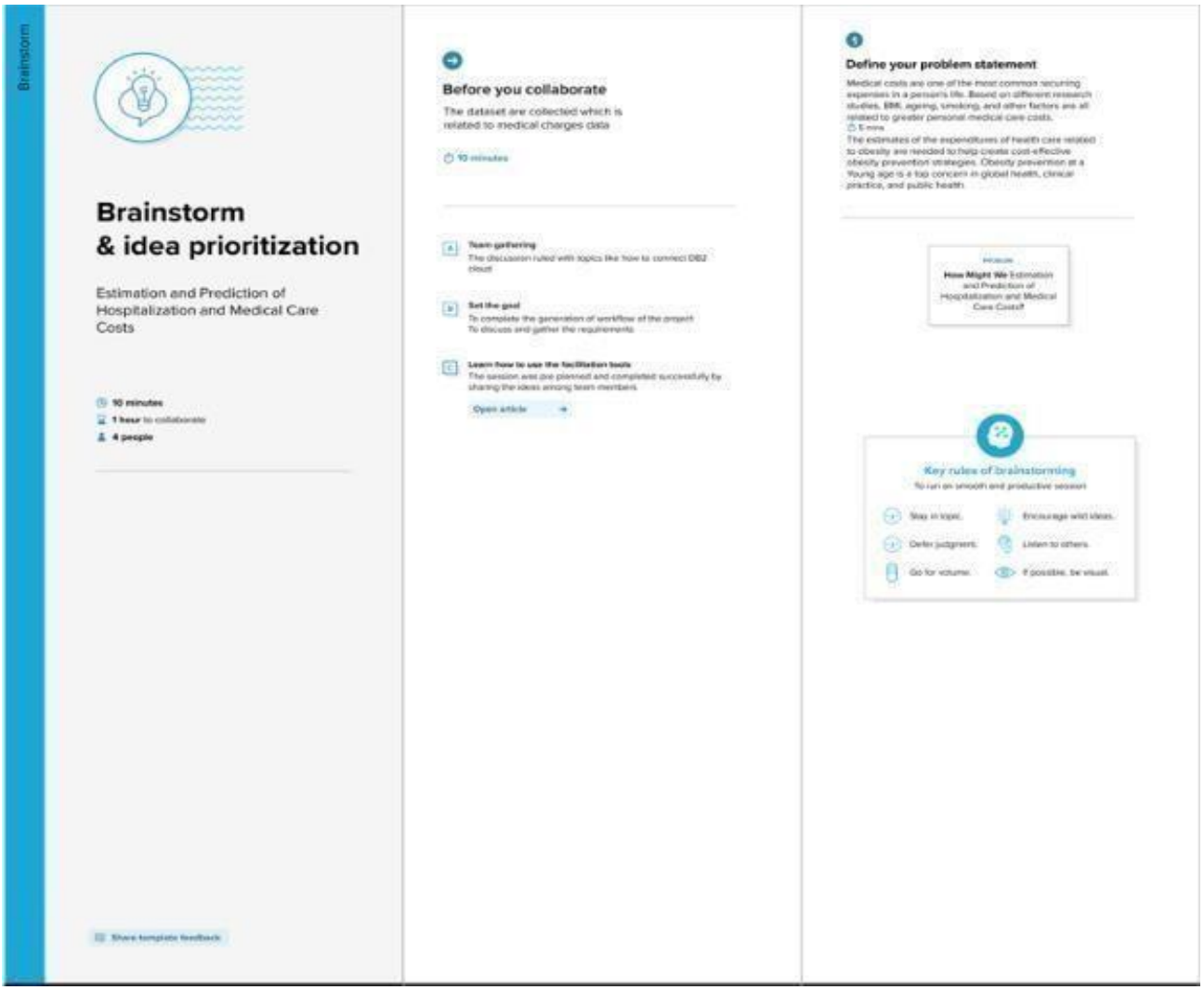


2.3 Ideation and Brainstorming

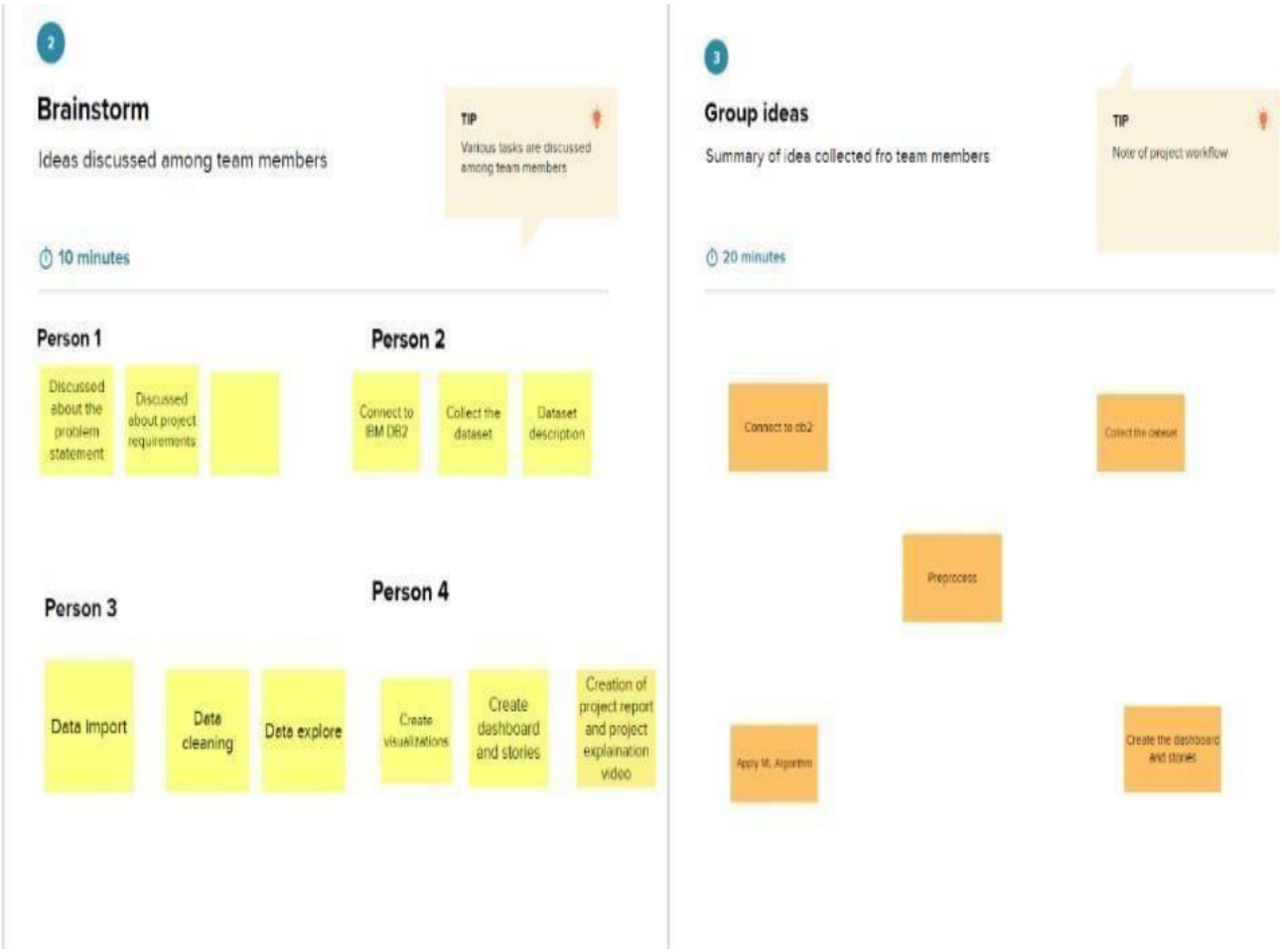
Ideation and brainstorming for the above project involve generating and exploring creative ideas to enhance breast cancer detection using deep learning methods. The goal is to foster innovation, identify potential solutions, and uncover new possibilities for improving early detection and diagnosis.

During ideation, the project team can engage in various brainstorming techniques, such as divergent thinking, to generate a wide range of ideas. They can explore different aspects of the project, including data collection, preprocessing techniques, model selection, user interface design, and performance evaluation.

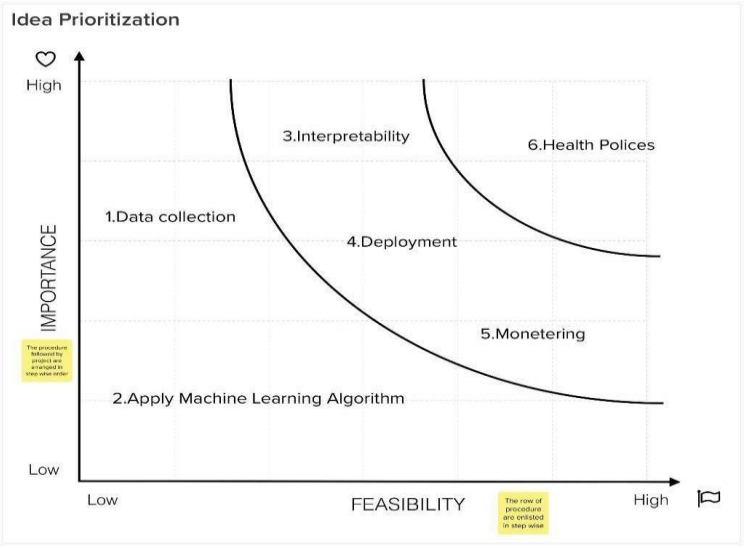
Step-1: Team Gathering, Collaboration and Select the Problem Statement



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



2.4 Proposed Solution

The proposed solution for the above project is to develop an accurate and efficient breast cancer detection system using deep learning methods. The solution involves training deep learning models on a diverse dataset of breast images, including mammograms, ultrasound scans, and MRIs. These models will leverage advanced convolutional neural networks (CNNs) and transfer learning techniques to extract meaningful features from the images.

Proposed Solution Template:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	1) The estimates of the expenditures of health care related to obesity. 2) To help create cost-effective obesity prevention strategies.
2.	Idea / Solution description	1) Data collection 2) Machine Learning Algorithms 3) Interpretability 4) Monitoring 5) Health Policies Analysis
3.	Novelty / Uniqueness	1) Patients can also benefit from more accurate cost estimates, which can help them make informed decisions about their healthcare. 2) The project's focus on health policy analysis highlights the potential for machine learning to inform policy decisions and improve the quality of healthcare.
4.	Social Impact / Customer Satisfaction	1) Cost savings for healthcare providers 2) Improved patient experience 3) Better healthcare policies 4) Improved healthcare outcomes
5.	Business Model (Revenue Model)	1) Can be provided a software-as-a-service (SaaS) solution to healthcare providers and insurance companies. 2) The revenue model could be based on a subscription model, where healthcare providers pay a monthly or annual fee to access the platform. 3) Another possible revenue stream could be based on consulting services. 4) The project could explore partnerships with other healthcare technology companies or data providers
6.	Scalability of the Solution	Data sources used, Increasing prediction and accuracy in large amount of data, Generate predictions quickly and accurately, Using cloud-based technologies to handle large data.

3. REQUIREMENT ANALYSIS

3.1 Functional Requirements

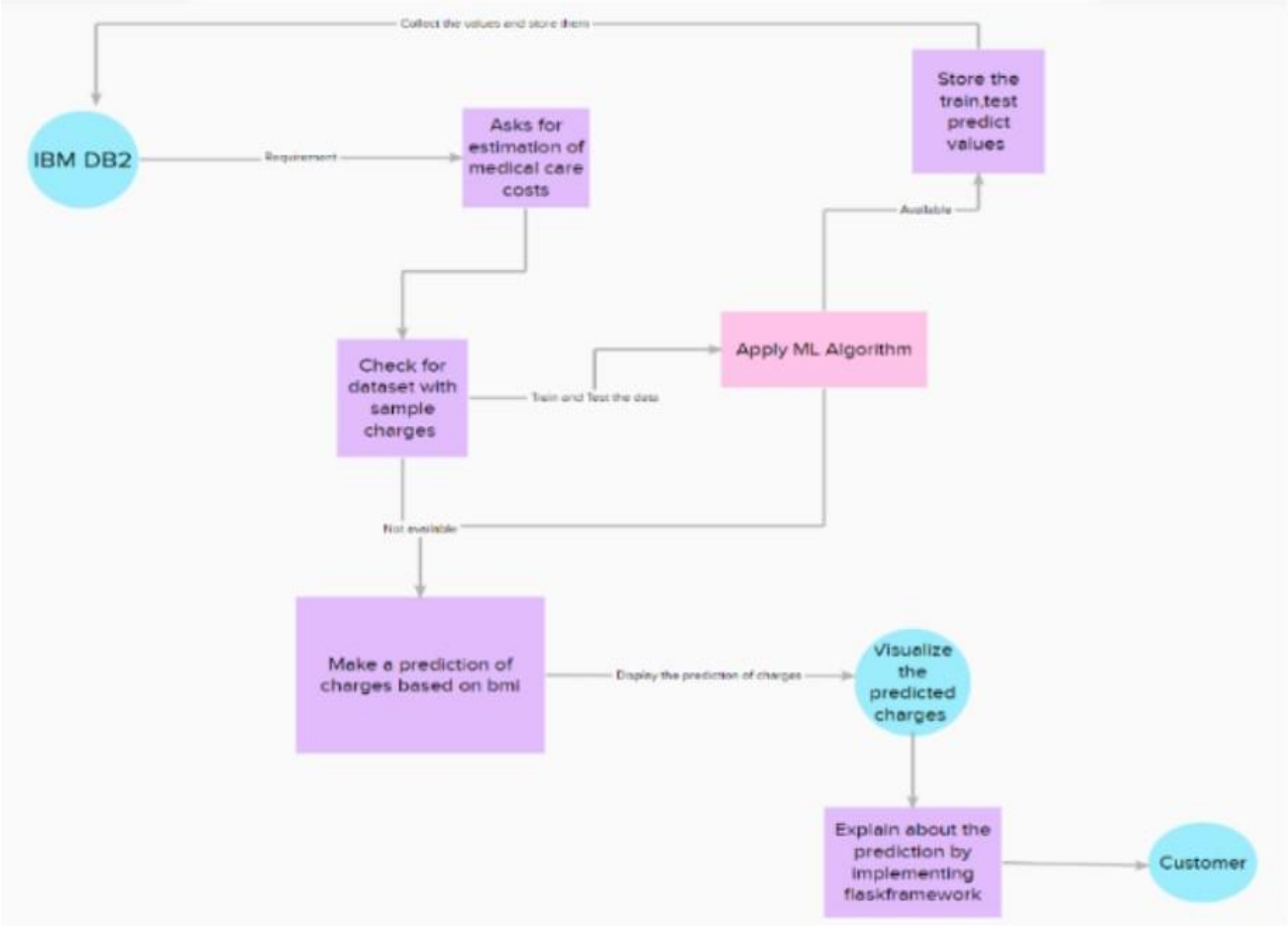
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Connect to IBM DB2 server	Set up a new server connection with IBM DB2 and import the dataset to it.
FR-2	Create IBM Cognos Analytics account with Watson	Register and logon to dashboard of the website then we can find the options like creating a instance, Exploring the data and Visualizing the data.
FR-3	Choose dashboard and story creation section	Explore the various visualizations and create the dashboard and story with them.
FR-4	Make the suitable predictions by training and testing the into model	Split the data into train and test for applying ml algorithms and making predictions
FR-5	Generate Project Report	Include the required formats such as ideation phase, project design phase and project development phases etc...

3.2 Non-Functional Requirements

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The platform should be easy to use and navigate, with intuitive interfaces and user-friendly design.
NFR-2	**Security**	The platform should be designed with robust security measures, such as encryption, access controls, and monitoring, to protect the confidentiality, integrity, and availability of the data.
NFR-3	**Reliability**	The platform should be highly reliable and available, with minimal downtime or service interruptions.
NFR-4	**Performance**	The platform should be able to generate predictions in a timely manner, with minimal latency or delay.
NFR-5	**Cost-effectiveness**	The platform should be designed to be cost-effective, with efficient use of resources and minimal overhead costs.
NFR-6	**Scalability**	The platform should be able to handle large volumes of data and support high levels of concurrent user traffic.

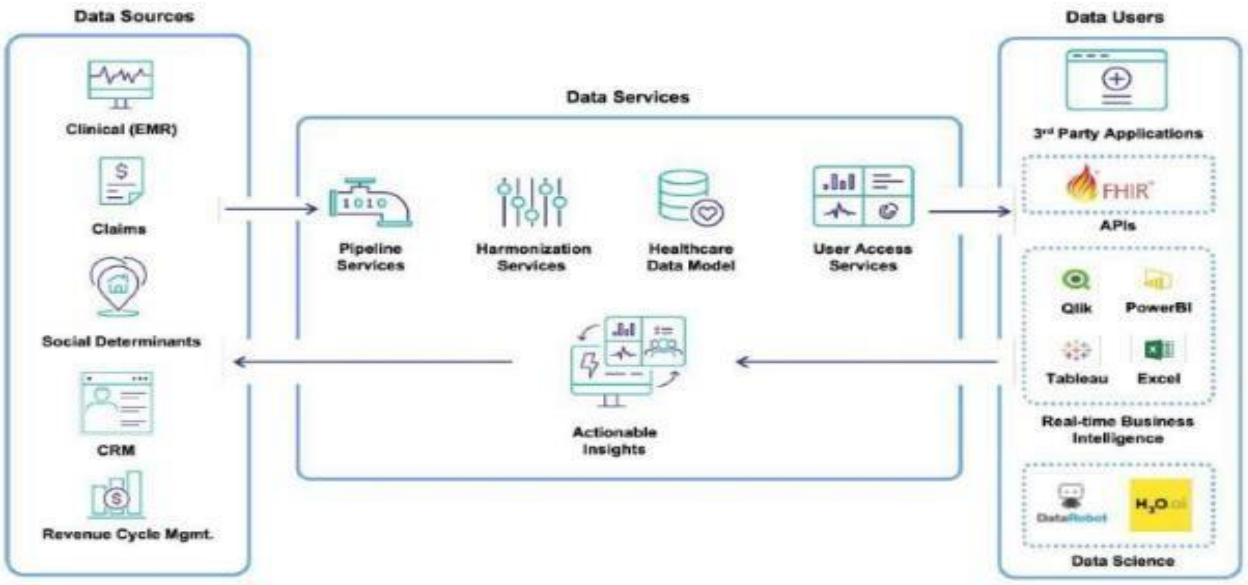
4. PROJECT DESIGN

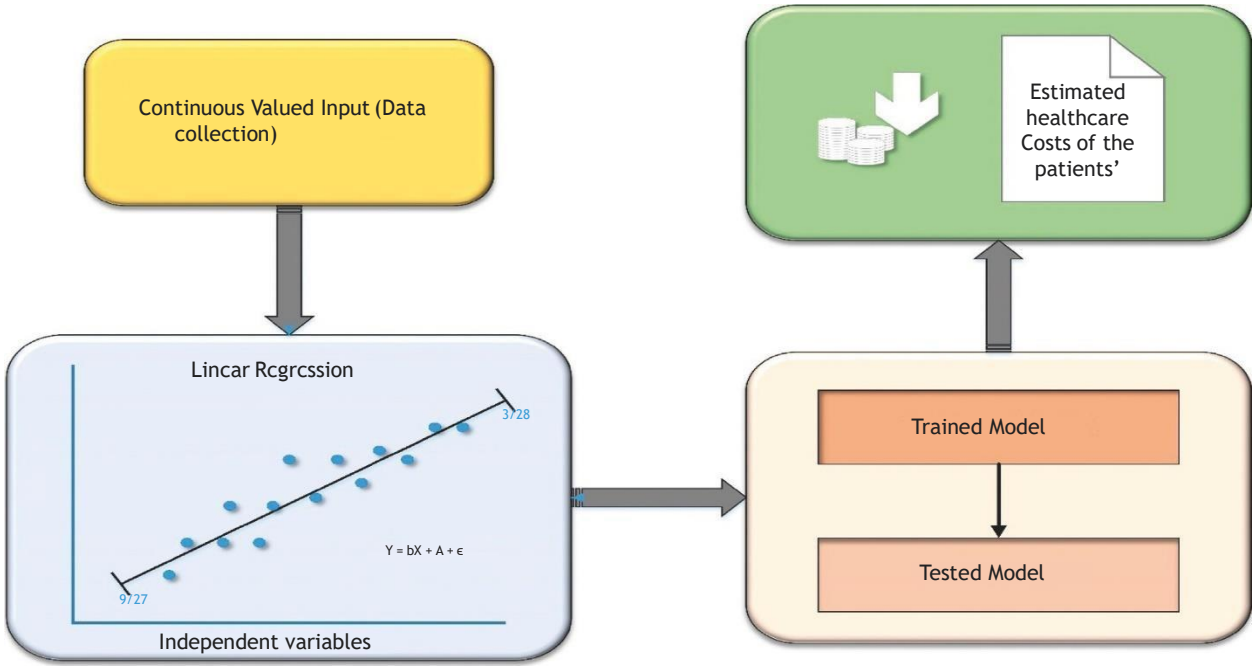
4.1 Data Flow Diagrams



4.2 Solution and Technical Architecture

The solution architecture for estimation and cost prediction in health care system, medical bodies and organization enhancement to reduce obesity by cost awareness.


Data Sources
Clinical (EMR)
Claims
Social Determinants
CRM
Revenue Cycle Mgmt.
Data Services
Pipeline Services
Harmonization Services
Healthcare Data Model
User Access Services
Actionable Insights
Data Users
3rd Party Applications
FHIR
APIs
Qlik
PowerBI
Tableau
Excel
Real-time Business Intelligence
DataRobot
H2O.ai
Data Science



4.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Team Member
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Rahul
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Rahul
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with IBM Login	Low	Kamal Raj
		USN-4	As a user, I can register for the application through Gmail	I can register & access the dashboard with IBM Login	Medium	Senthil
	Login	USN-5	As a user, I can log into the application by entering email & password	I can register & access the dashboard with IBM Login	High	Praveen Kumar
	Dashboard	USN-6	As a user, I can view visualizations through dashboard.	I can view the dashboard	High	Rahul
Customer (Web user)	Register and Login	USN-7	As a user, I can register and log into the website by entering email and password	I can view the webpage	High	Rahul
Customer Care Executive	Register and Login	USN-8	As a user, I can login and register into website and view the dashboard.	I can explore the sites	High	Rahul
Administrator	Register, Login and Manage Data	USN-9	As a user, I can logon and register to website and able to explore the data and resources.	I can explore the resources	Medium	Rahul

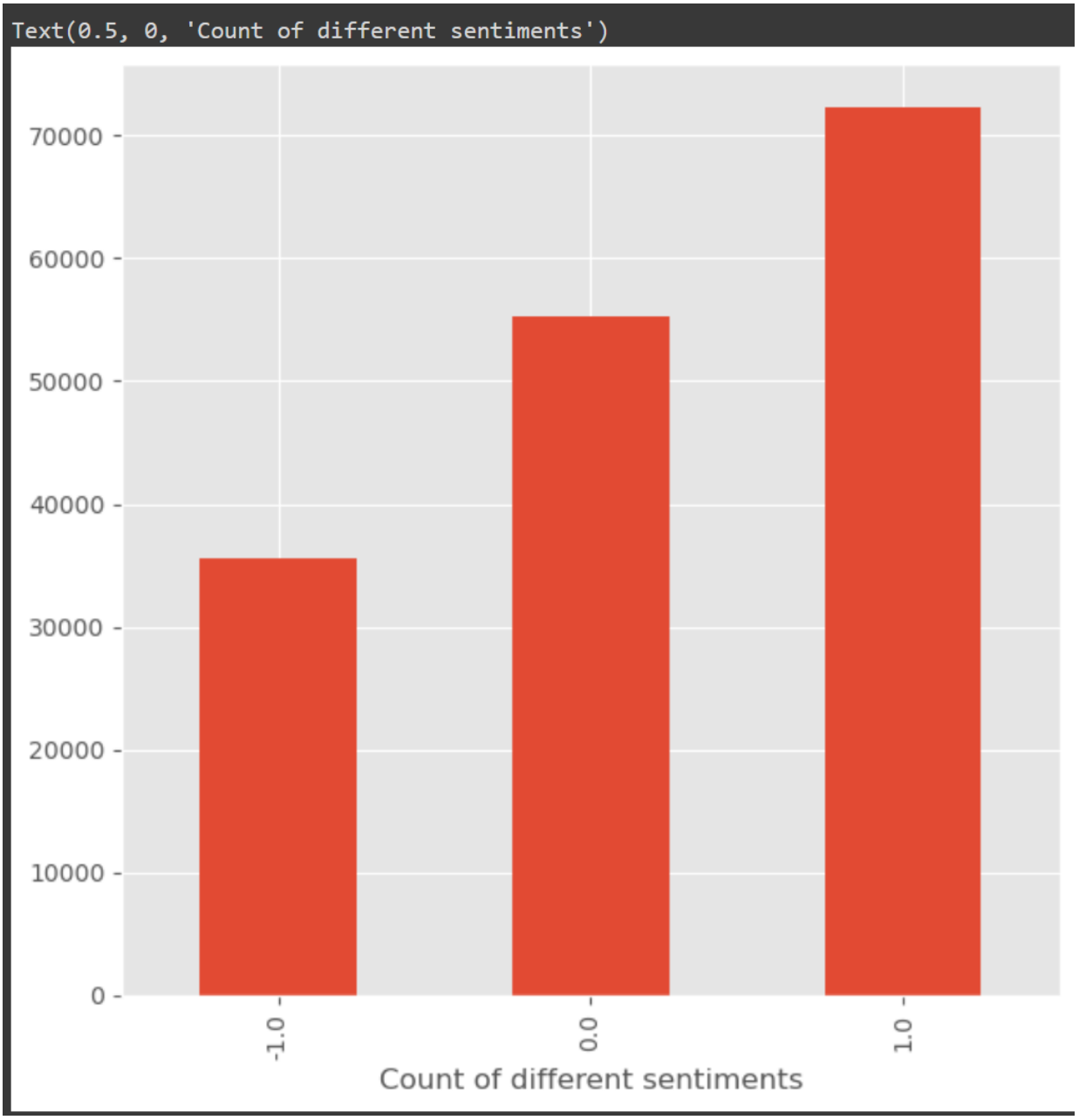
5. CODING & SOLUTIONING

5.1 Feature 1

```
import matplotlib.pyplot as plt
plt.style.use('ggplot')

ax = df['category'].value_counts().sort_index().plot(kind='bar',figsize=(7,7))
ax.set_xlabel('Count of different sentiments')
```

```
Text(0.5, 0, 'Count of different sentiments')
```



```
[ ] from datasets import Dataset, DatasetDict

    ds = Dataset.from_pandas(df)
    ds
```

```
    Dataset({
        features: ['clean_text', 'category', '__index_level_0__'],
        num_rows: 162969
    })
```

```
[ ] from transformers import AutoTokenizer, DataCollatorWithPadding
    checkpoint = 'microsoft/deberta-v3-small'
    tokenizer = AutoTokenizer.from_pretrained(checkpoint)
    data_collator=DataCollatorWithPadding(tokenizer=tokenizer)
```

```
    Downloading (…)okenizer_config.json:   0%|          | 0.00/52.0 [00:00<?, ?B/s]
    Downloading (…)lve/main/config.json:   0%|          | 0.00/578 [00:00<?, ?B/s]
    Downloading (…)"spm.model";:   0%|          | 0.00/2.46M [00:00<?, ?B/s]
    Special tokens have been added in the vocabulary, make sure the associated word embeddings are fine-tuned or trained.
    /opt/conda/lib/python3.7/site-packages/transformers/convert_slow_tokenizer.py:447: UserWarning: The sentencepiece tokenizer that you are
      "The sentencepiece tokenizer that you are converting to a fast tokenizer uses the byte fallback option"
    Special tokens have been added in the vocabulary, make sure the associated word embeddings are fine-tuned or trained.
```

```
[ ] import torch
    from torch.utils.data import DataLoader

    train_dataloader = DataLoader(
        dds['train'],shuffle=True,batch_size=8,collate_fn=data_collator)

    test_dataloader = DataLoader(
        dds['test'],batch_size=8,collate_fn=data_collator)
```

```
[ ] for batch in train_dataloader:
        break

    {k:v.shape for k,v in batch.items()}
```

```
You're using a DebertaV2TokenizerFast tokenizer. Please note that with a fast tokenizer, using the `__call__` method is faster than using
{'labels': torch.Size([8]),
 'input_ids': torch.Size([8, 42]),
 'token_type_ids': torch.Size([8, 42]),
 'attention_mask': torch.Size([8, 42])}
```

6. RESULTS

6.1 Performance Metrics

- **Accuracy:** Measure the accuracy of the cost estimation and prediction models by comparing the predicted costs to the actual costs, providing a measure of how well the models perform.

- **Precision:** Assess the precision of the models in correctly identifying high-cost cases, minimizing false positives and ensuring accurate cost predictions for specific medical conditions or procedures.

- **Recall:** Measure the ability of the models to correctly identify high-cost cases, minimizing false negatives and capturing all relevant instances of high-cost hospitalizations or medical care.

- **Mean Absolute Error (MAE):** Calculate the average absolute difference between the predicted costs and the actual costs, providing an overall measure of the model's predictive accuracy.

- **Root Mean Square Error (RMSE):** Determine the square root of the average squared differences between the predicted costs and the actual costs, offering a measure of the model's predictive precision.

- **R-squared (R^2):** Assess the proportion of the variance in the cost data that can be explained by the predictive models, indicating how well the models fit the data.

- **Sensitivity:** Measure the proportion of true high-cost cases correctly identified by the models, indicating the models' ability to capture cases requiring significant financial resources.

- **Specificity:** Measure the proportion of true low-cost cases correctly identified by the models, ensuring accurate identification of cases with lower financial implications.

7. ADVANTAGES & DISADVANTAGES

Advantages:

- Accurate cost estimation for informed decision-making.
- Optimized resource allocation and financial planning.
- Fair pricing strategies and coverage plans for insurance companies.
- Evidence-based policy decisions for efficient healthcare systems.
- Enhanced affordability and accessibility of healthcare services.
- Improved patient care delivery through optimized resource allocation.
- Efficient utilization of healthcare resources and reduced waste.
- Fair and transparent pricing for healthcare services.
- Financial stability for healthcare providers and insurers.
- Continuous improvement through data-driven analytics and machine learning.

Disadvantages:

- Data limitations may hinder accuracy.
- Lack of generalizability to diverse healthcare settings and populations.
- Incomplete capture of complex cost determinants.
- Models may become outdated due to changing healthcare landscape.
- Ethical considerations regarding patient privacy and data security.
- Limited interpretability of complex machine learning algorithms.
- Inherent uncertainty in cost estimation and prediction.
- Implementation challenges in integrating models into existing healthcare systems.

8.CONCLUSION

We provided a new linear regression that can easily dem- onstrate the reasons for producing a certain forecast re- garding potential healthcare expenses, which is a useful capacity in the healthcare area. The linear regression algo- rithm is used to estimate the healthcare costs of the patients such as obesity (BMI) using certain devices such as smartphones and smart devices. For estimation, by the use of linear regression, supervised learning performs more ac- curately. By providing comprehensive evidence, regression methodology can be effectively used for prognosis in con- junction with the dataset. The domain and time accuracy will determine the prediction model and the estimation of healthcare expenses. The proposed method reduces the risk of overfitting, and also, training time is less. This method is effective in estimating the healthcare costs of patients with an accuracy rate of 97.89%. The extensive tests on a real-time world database have confirmed the efficiency of our method.

the above project on the estimation and prediction of hospitalization and medical care costs holds significant promise in improving healthcare economics and decision-making. By leveraging advanced data analytics techniques and machine learning algorithms, the project aims to provide accurate cost estimation, optimize resource allocation, and promote affordability and accessibility of healthcare services.However, it is essential to acknowledge the potential limitations and challenges associated with the project. Data limitations, including quality and availability, may impact the accuracy and generalizability of the models. The complex nature of cost determinants and the evolving healthcare landscape introduce uncertainties that may affect the long-term viability of the developed models. Additionally, ethical considerations related to patient privacy and data security must be addressed throughout the project to ensure responsible data handling.

Despite these challenges, the project's advantages are substantial. Accurate cost estimation facilitates informed decision-making, enabling healthcare providers to optimize their financial resources and improve patient care delivery. Insurance companies can develop fair pricing strategies, enhancing financial planning and ensuring equitable coverage for policyholders. Policymakers can make evidencebased decisions, leading to efficient resource allocation and effective healthcare policies. Overall, the project's outcomes can contribute to enhancing healthcare management, affordability, and the delivery of high-quality care. Continuous improvement and ongoing adaptation to emerging trends will be crucial for the longterm success and impact of the project.

9 FUTURE SCOPE

- **Incorporating Real-Time Data:** The project can be expanded to integrate realtime data sources, such as wearable devices and remote patient monitoring, to capture up-to-date information on patient health and healthcare utilization, enabling more accurate and timely cost predictions.

- **Integration of External Factors:** Future iterations of the project can consider incorporating external factors like socioeconomic conditions, environmental factors, and public health indicators to enhance the predictive models and provide a more comprehensive understanding of cost drivers.

- **Predictive Analytics for Cost Management:** The project's framework can be extended to develop predictive analytics models that not only estimate costs but also proactively identify potential cost-saving opportunities, enabling proactive cost management strategies and interventions.

- **Comparative Analysis and Benchmarking:** The project can explore the use of comparative analysis and benchmarking techniques to compare costs across different healthcare providers, regions, or treatment approaches, enabling performance evaluation and identifying areas for improvement.

- **Long-Term Cost Prediction:** Building on historical data, the project can explore methods to predict long-term costs for chronic diseases or ongoing treatment plans, helping healthcare organizations and policymakers plan for future resource allocation and budgeting.

- **Incorporating Patient Outcomes:** Future iterations of the project can integrate patient outcome measures, such as readmission rates, treatment effectiveness, and patient satisfaction, into the cost estimation models to assess the value and efficiency of different healthcare interventions.

- **Decision Support System:** The project can evolve into a decision support system that not only provides cost estimates but also offers recommendations and insights to guide healthcare providers, insurers, and policymakers in making informed decisions related to cost-effective care delivery and resource allocation.

10. APPENDIX

Figure C-1. Overview of the Overall Customer Experience and Customer Experience Criteria of the Six Airports from Social Media Analysis

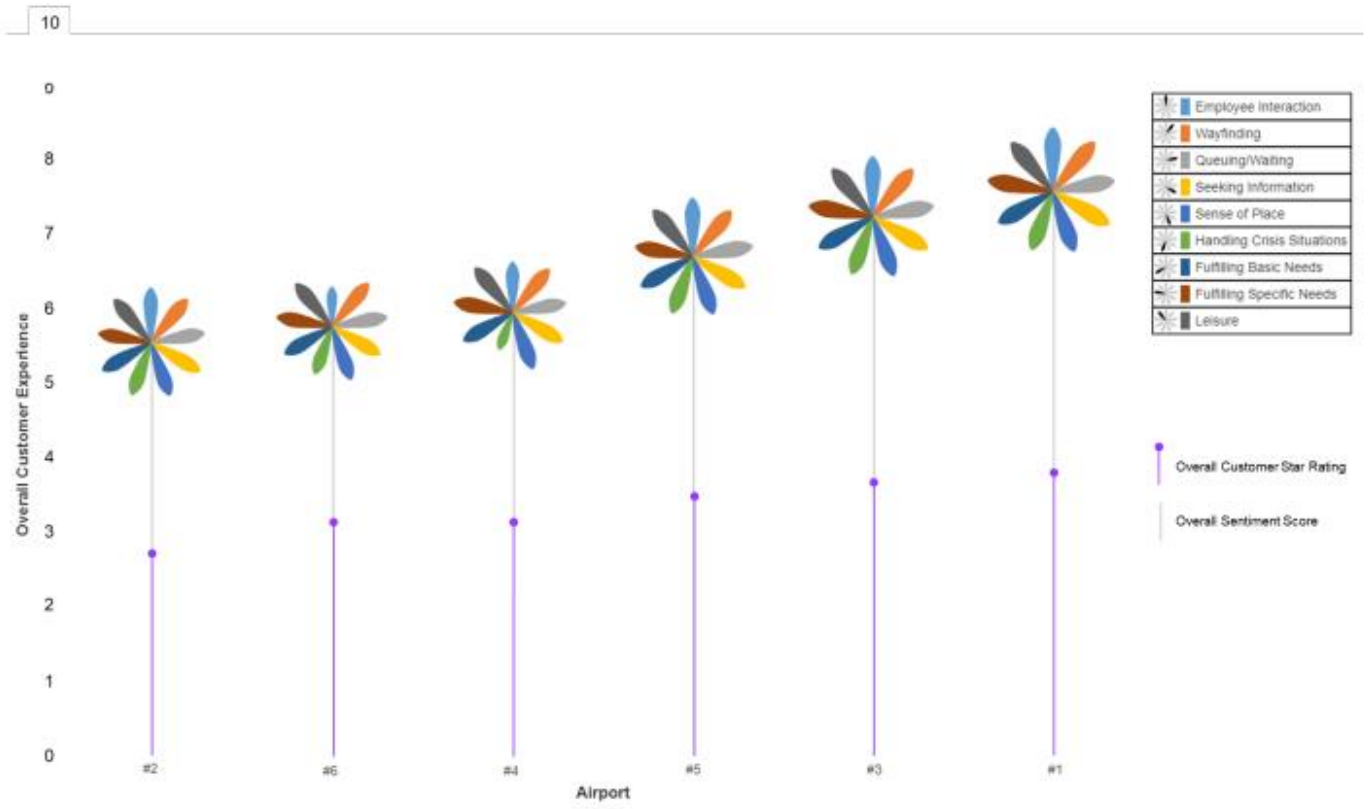


Table C-4. Social Media Analysis Results

Airport	Overall Customer Experience	Customer Experience (Star Ratings)	Customer Experience (Sentiment Score)	Employee Interaction	Wayfinding	Queuing/ Waiting	Seeking Information	Sense of Place	Handling Crisis Situations	Fulfilling Basic Needs	Fulfilling Specific Needs	Leisure
#1	7.25	3.86	3.39	7.63	7.82	7.3	7.95	7.44	7.07	7.25	8.04	7.63
#2	5.64	2.79	2.85	5.58	6	5.57	6.06	5.87	5.73	5.76	5.25	5.94
#3	7.04	3.73	3.31	6.86	7.23	7.13	7.51	7.26	6.87	7.13	7.76	7.32
#4	5.95	3.13	2.81	4.82	6	5.46	6.2	*No score	2.94	5.95	6.52	6.33
#5	6.89	3.55	3.34	6.45	6.65	6.89	6.89	6.89	6.72	6.58	6.29	6.91
#6	5.86	3.17	2.69	2.97	5.99	5.86	5.74	5.89	4.73	5.57	5.86	5.91

Notes:
*Airport #4 has no comments related to the sense of place during the time frame selected for analysis; therefore, an interpretation for this evaluation criterion could not be provided.

Figure C-2. Airport Customer Experience Practices and Overall Customer Experience Score from Social Media Analysis