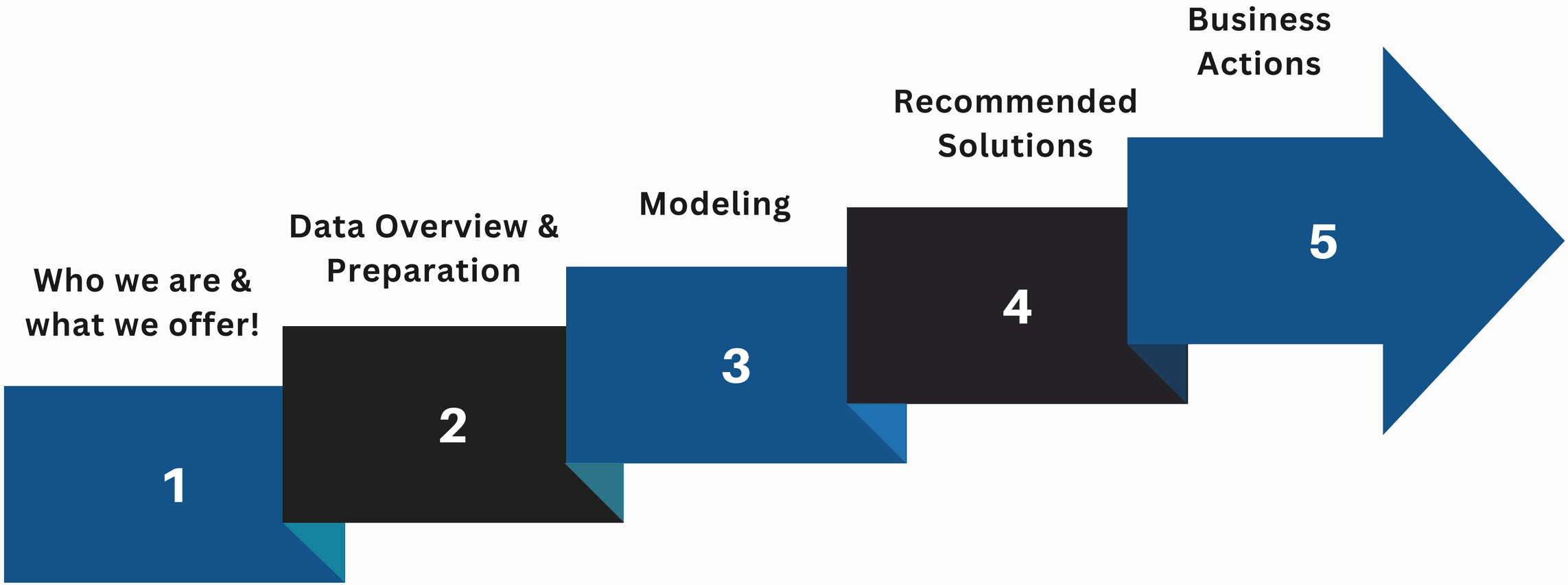
YelpSense

Unleashing the Power of Customer Feedback

Presented by:

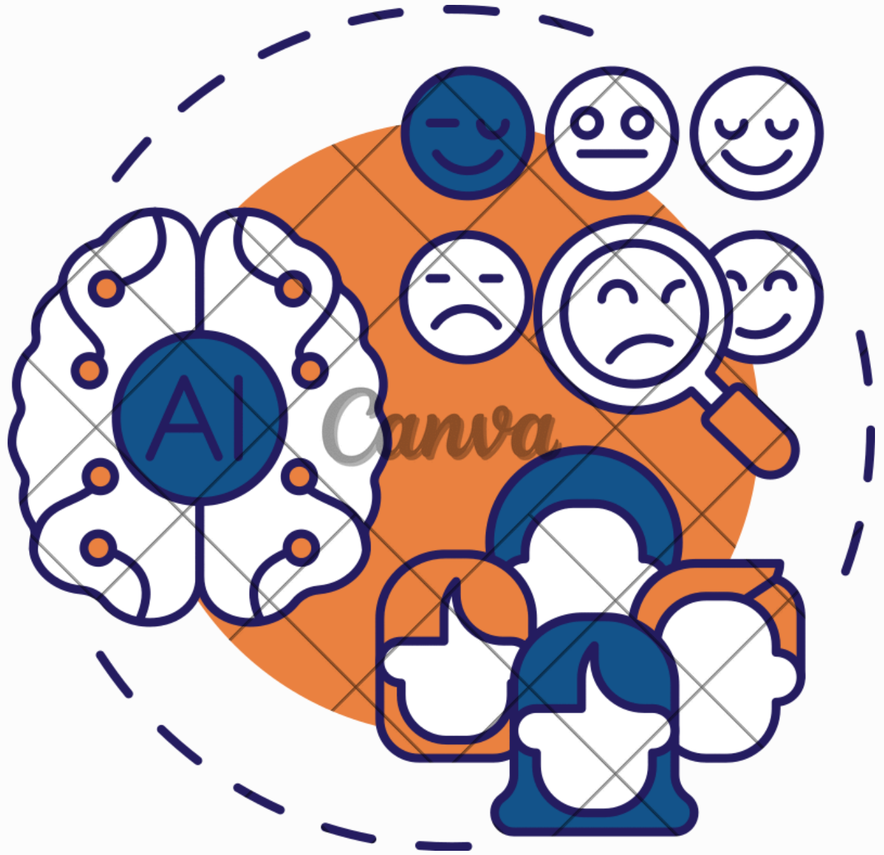
Anushka Shah, Yashi Tiwari, Anant Bairagi,
Naga 'Nitesh' Sanka, Haridakshini 'Harisha' Ajeetha

Agenda


Who we are & what we offer!
1
Data Overview & Preparation
2
Modeling
3
Recommended Solutions
4
Business Actions
5

Business Problem Statement

60% of people read reviews before visiting restaurant. With the rise of online reviews, businesses face the challenge of analyzing vast amounts of customer feedback to understand their sentiment and make data-driven decisions.Traditional manual methods are time-consuming, subjective, and not scalable to handle big data

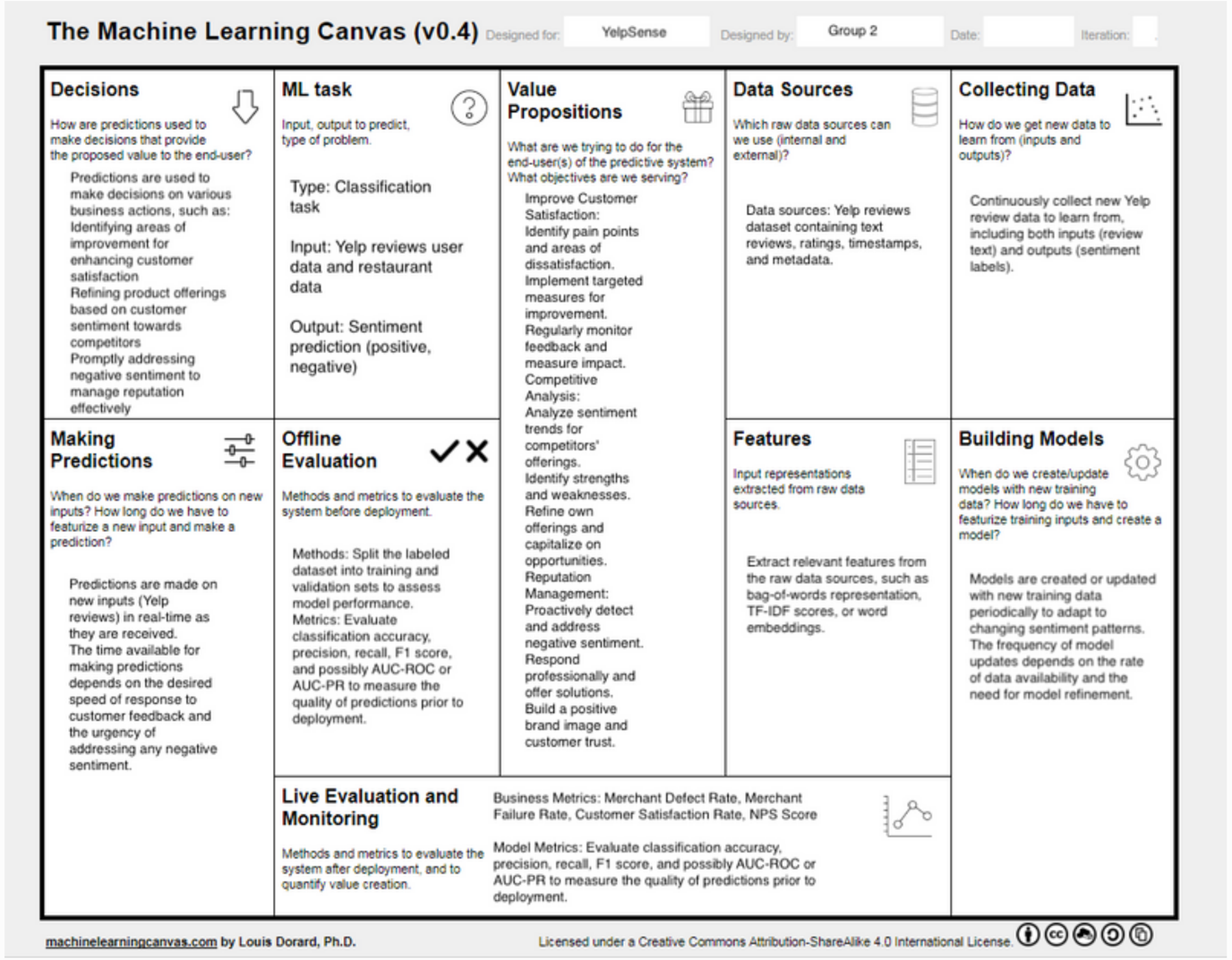


Who we are?

Unleashing the Power of Customer Feedback

YelpSense is a leading provider of advanced sentiment analysis solutions for product reviews. We aim to provide an automated, scalable, and accurate sentiment analysis solution to help businesses extract valuable insights from Yelp reviews.

Machine Learning Canvas



The Machine Learning Canvas (v0.4)

Designed for: YelpSense | Designed by: Group 2 | Date: | Iteration:

Decisions

How are predictions used to make decisions that provide the proposed value to the end-user?

Predictions are used to make decisions on various business actions, such as:
Identifying areas of improvement for enhancing customer satisfaction
Refining product offerings based on customer sentiment towards competitors
Promptly addressing negative sentiment to manage reputation effectively

ML task

Input, output to predict, type of problem.

Type: Classification task

Input: Yelp reviews user data and restaurant data

Output: Sentiment prediction (positive, negative)

Value Propositions

What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

Improve Customer Satisfaction:
Identify pain points and areas of dissatisfaction.
Implement targeted measures for improvement.
Regularly monitor feedback and measure impact.
Competitive Analysis:
Analyze sentiment trends for competitors' offerings.
Identify strengths and weaknesses.
Refine own offerings and capitalize on opportunities.
Reputation Management:
Proactively detect and address negative sentiment.
Respond professionally and offer solutions.
Build a positive brand image and customer trust.

Data Sources

Which raw data sources can we use (internal and external)?

Data sources: Yelp reviews dataset containing text reviews, ratings, timestamps, and metadata.

Collecting Data

How do we get new data to learn from (inputs and outputs)?

Continuously collect new Yelp review data to learn from, including both inputs (review text) and outputs (sentiment labels).

Making Predictions

When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?

Predictions are made on new inputs (Yelp reviews) in real-time as they are received.
The time available for making predictions depends on the desired speed of response to customer feedback and the urgency of addressing any negative sentiment.

Offline Evaluation

Methods and metrics to evaluate the system before deployment.

Methods: Split the labeled dataset into training and validation sets to assess model performance.
Metrics: Evaluate classification accuracy, precision, recall, F1 score, and possibly AUC-ROC or AUC-PR to measure the quality of predictions prior to deployment.

Features

Input representations extracted from raw data sources.

Extract relevant features from the raw data sources, such as bag-of-words representation, TF-IDF scores, or word embeddings.

Building Models

When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?

Models are created or updated with new training data periodically to adapt to changing sentiment patterns.
The frequency of model updates depends on the rate of data availability and the need for model refinement.

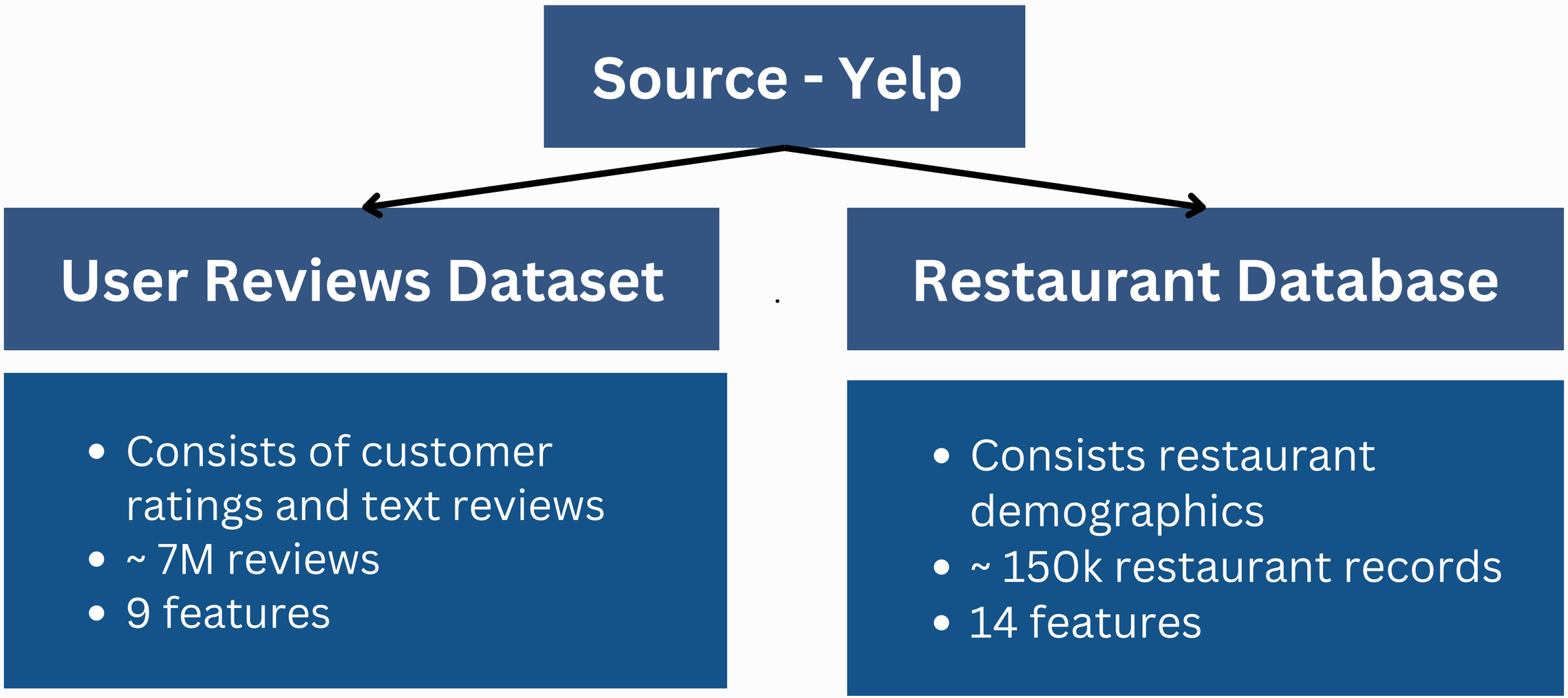
Live Evaluation and Monitoring

Methods and metrics to evaluate the system after deployment, and to quantify value creation.

Business Metrics: Merchant Defect Rate, Merchant Failure Rate, Customer Satisfaction Rate, NPS Score

Model Metrics: Evaluate classification accuracy, precision, recall, F1 score, and possibly AUC-ROC or AUC-PR to measure the quality of predictions prior to deployment.

Data Overview



Data Preparation

Review Dataset - Selected 3 features

- Business ID
- Review Stars
- Review Text

Restaurant Dataset - Selected 2 features

- Business ID
- Category of Restaurant

Inner join on Business ID

- Analyzed "English Reviews **ONLY**"
- Removed 3-star reviews
- Labelling - Sentiment
  - Stars < 3 - Classified as 0
  - Stars > 3 - Classified as 1

FINAL DATASET

- **Balanced 0 and 1 reviews**
  - **(21.3% vs 78.6%)**
- **Removed stop words**
- **Final columns:**
  - **Cleaned reviews &**
  - **Sentiment**

Model Building

1	Data Randomly Split : 70:30
2	Feature Engineering and Text Vectorization Models • Bag-of-Words Representation • TF-IDF • nGrams
3	Classification Models • Logistic Regression • Random Forest • SVM
4	Model Evaluation and Integration

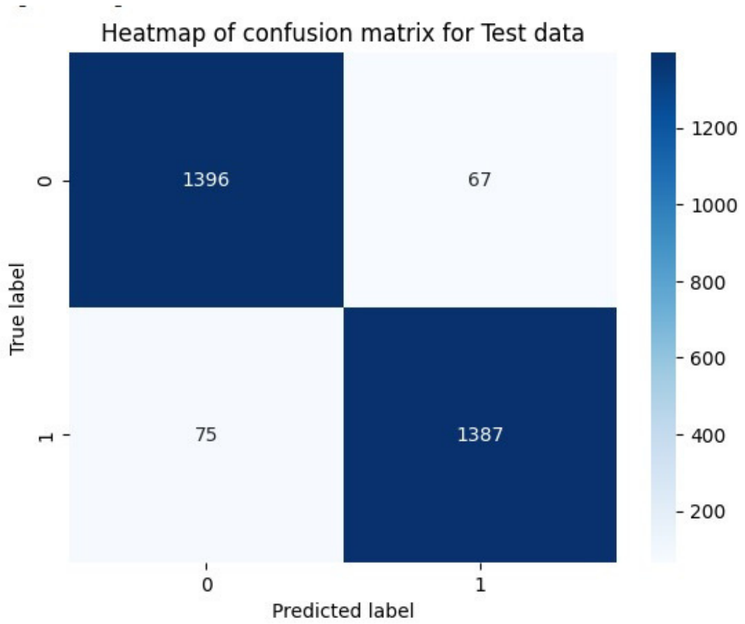
Model Evaluation

Using TD-IF

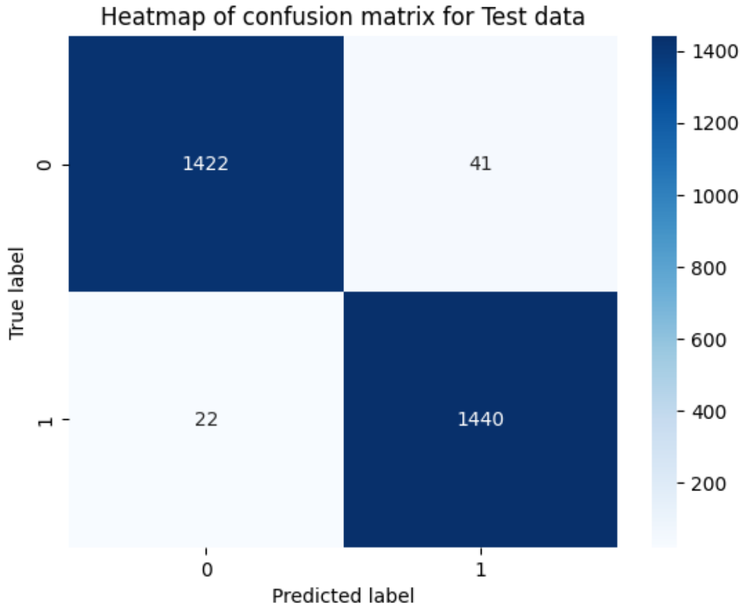
Models Metrics	Logistic	Random Forest	SVM
Accuracy	0.95	0.94	**0.97**
Precision	0.95	0.97	**0.97**
Recall	0.94	0.97	**0.98**
F1 Score	0.95	0.97	**0.97**

Confusion Matrix

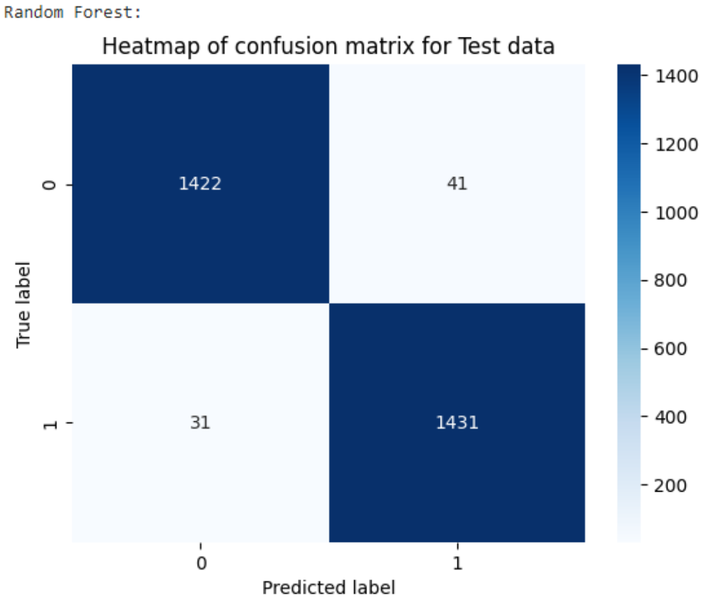
Logistics Regression



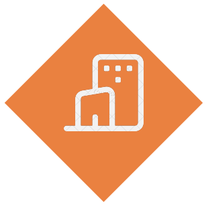
SVM



Random Forest

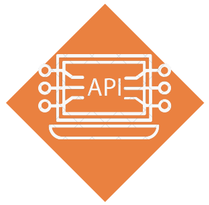


Integration and Deployment



Seamless Integration

Our sentiment analysis solution can be seamlessly integrated into existing systems and workflows.



Customization Options

We offer customization options to align with specific business requirements and objectives ensuring a tailored approach.

Scalable Deployment

Our solution is designed to handle large volumes of Yelp reviews, making it scalable to accommodate growing data needs as the business expands

User-Friendly Interface

Our solution comes with a user-friendly interface that allows businesses to easily access and interpret sentiment analysis results, empowering decision-makers with actionable insights.

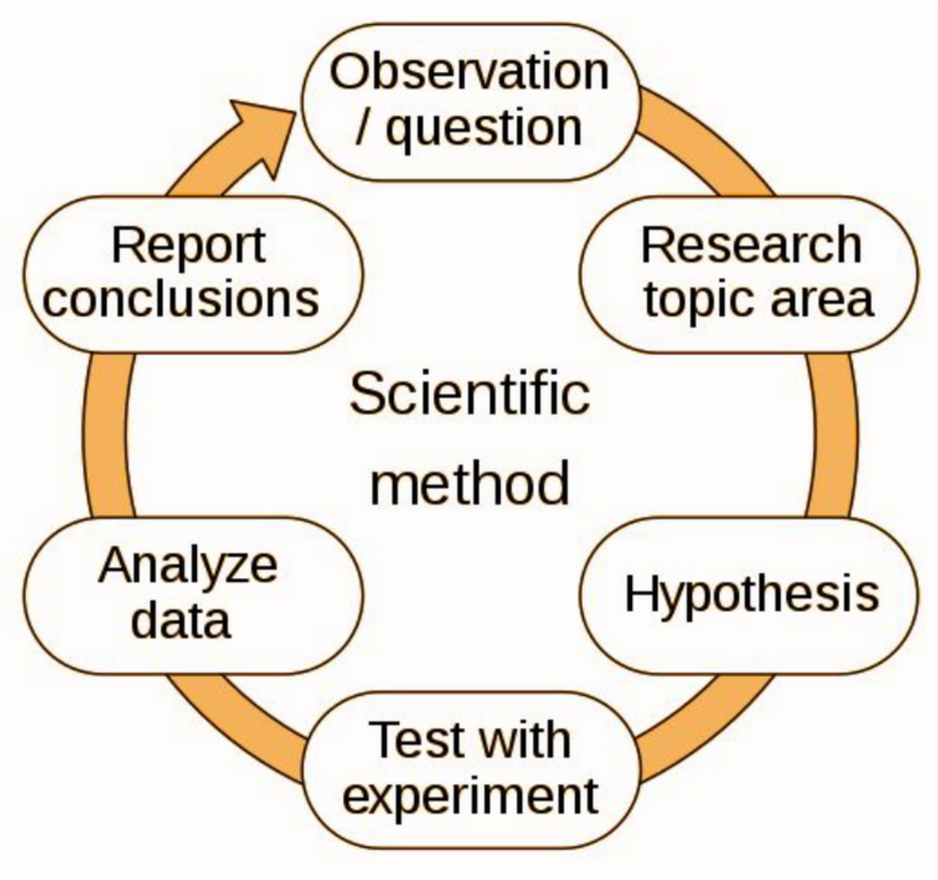
Recommended Business Solution

ADVANCED NLP	MACHINE LEARNING MODELS	REAL-TIME MONITORING AND ALERT SYSTEMS
• State-of-the-art algorithms for sentiment analysis • Contextual understanding and nuanced language analysis .	• Custom sentiment analysis models • Accurate classification of reviews (positive, negative, neutral)	• Continuous scanning of Yelp reviews • Prompt identification of negative sentiment

How can we help?

IMPROVE CUSTOMER SATISFACTION	COMPETITIVE EDGE	BRAND IMAGE
• Identify pain points and areas of dissatisfaction. • Implement targeted measures for improvement. • Regularly monitor feedback and measure impact.	• Analyze sentiment trends for competitors' offerings. • Identify strengths and weaknesses. • Refine own offerings and capitalize on opportunities.	• Proactively detect and address negative sentiment. • Respond professionally and offer solutions. • Build a positive brand image and customer trust.

Conclusion



We offer businesses a powerful solution by leveraging advanced NLP techniques, ML models, and real-time monitoring to extract valuable insights from customer feedback at scale, enabling data-driven decision-making.

Our solution empowers businesses to unlock the full potential of Yelp reviews and drive business outcomes through actionable sentiment analysis.

Thank You

Invest in our vision and join us on this exciting journey