

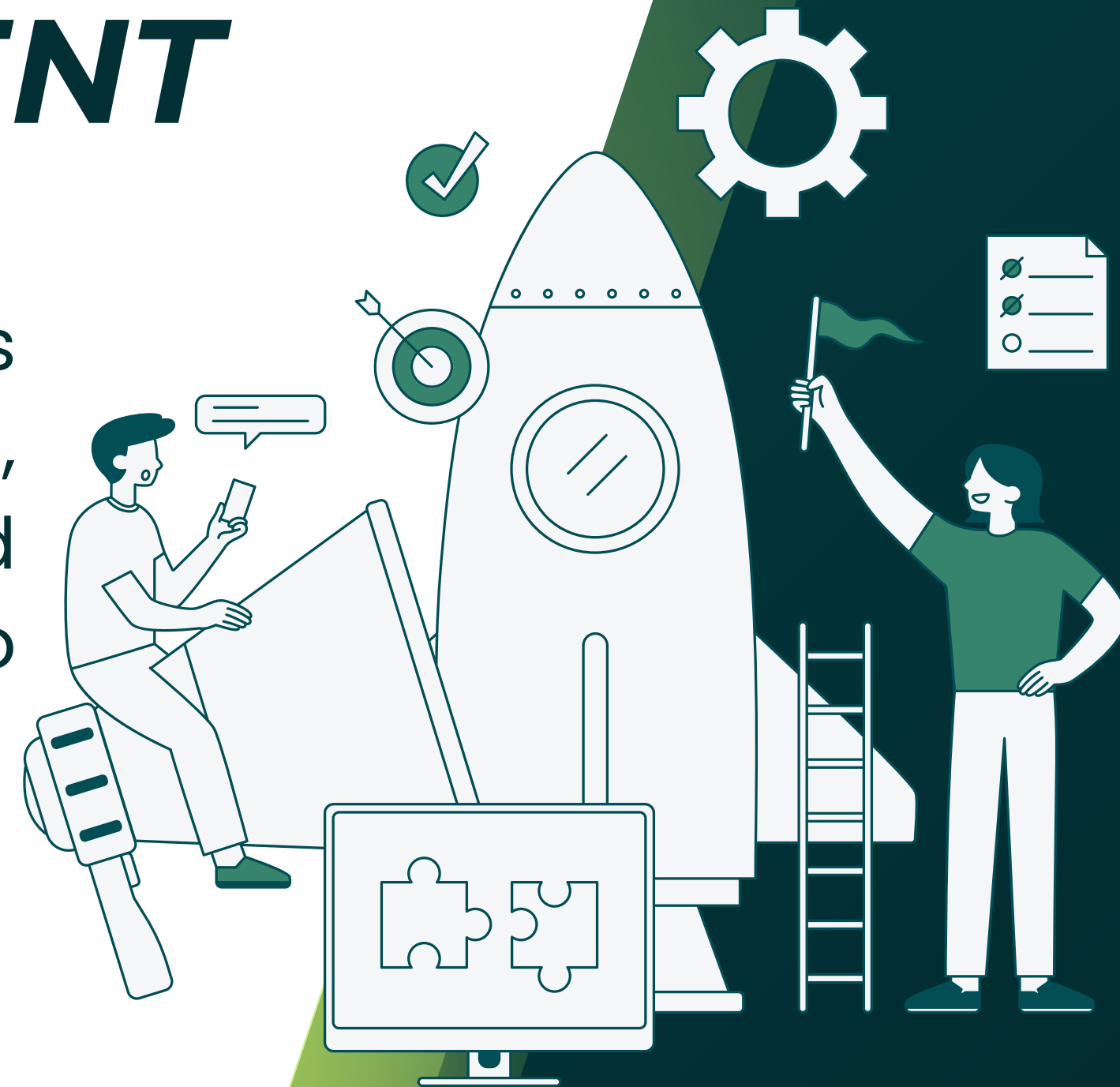
INTEL PROJECT

**PS-11: INTEL PRODUCTS SENTIMENTAL
ANALYSIS FROM ONLINE REVIEWS**

By: Ideafactory team

PROBLEM STATEMENT

This project aims to analyse user reviews of Intel product from an e-commerce site, focusing on scraping reviews and performing sentiment analysis to understand customer feedback.



IDEA BRIEFING

We have approached the solution by combining the strengths of automated web scraping and dual sentiment analysis models to extract and analyze customer reviews effectively.

For sentimental analysis, we employed 2 distinct models each tailored to specific scenarios :

BERT MODEL

VADER



FEATURES OFFERED

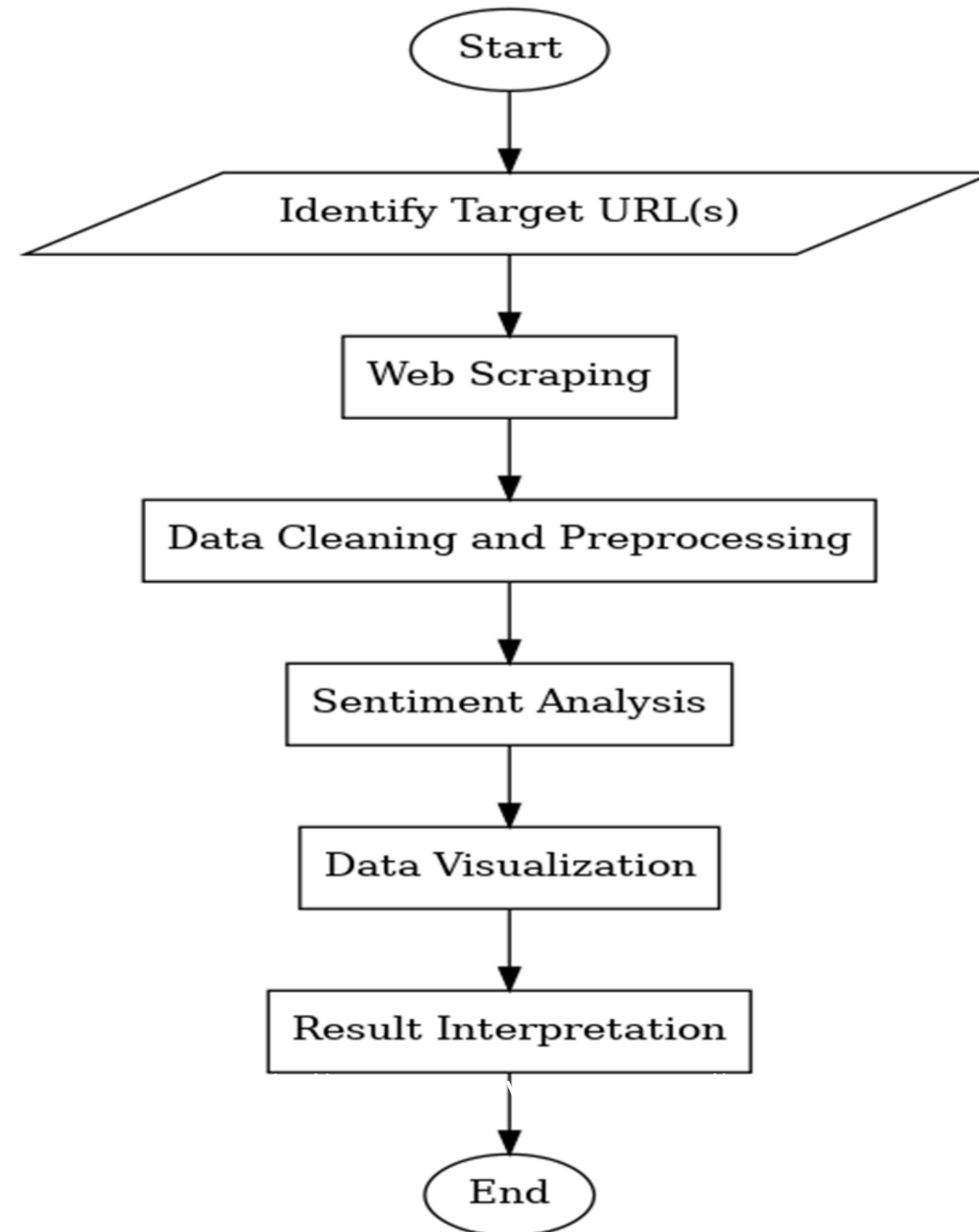
- **Web Scraping:** Gathering review data from various online sources.
- **Data Cleaning and Preprocessing:** Preparing text data for analysis by removing noise and standardizing formats.
- **Sentiment Analysis:** Using tools like VADER and BERT model to analyze and categorize sentiment.
- **Visualization:** Displaying sentiment distribution and other insights through graphs and chart
- **Feature Extraction:** Identifying and extracting relevant features from the text data, such as keywords or phrases.
- **Trend Analysis:** Tracking changes in sentiment over time to identify patterns or shifts in opinions.

TECHNOLOGIES USED

libraries that have been employed in the model :

- **Python**
- **BeautifulSoup** for web scraping
- **NLTK** for natural language processing
- **VADER** and **Text Blob** for sentiment analysis
- **BERT** For deep learning-based, context-rich sentiment classification.
- **Pandas** for data manipulation
- **Matplotlib** and **Seaborn** for visualization
- **Word Cloud** For generating visual representations of word frequency to highlight important terms in the reviews.

PROCESS FLOW



WEB SCRAPPING

ARCHITECTURE DESIGN

Data Source

Identifies the websites or online resources where data will be scraped.

Web Scraper

- The tool or software used to extract data from the identified sources.
- Uses techniques like HTML parsing, DOM traversal, or API calls to collect data.
- Ensures data is collected in a structured format suitable for further processing.

Data Storage:

- Stores the scraped data in a database or file system. Ensures the data is organized and accessible for analysis.

Sentiment Analysis

- Applies sentiment analysis techniques to the stored data.
- Output sentiment scores or categories based on the analysis.

BERT MODEL

ARCHITECTURE DESIGN

Text Input

Initial step where raw text data is fed into the system.

BERT Tokenizer

- Converts words into sub-words, characters, or word pieces suitable for BERT processing.
- Ensures that the tokenized text can be effectively understood by the BERT model.

BERT Model

- The core of the sentiment analysis architecture.
- Uses a pre-trained BERT model that understands the context of the text.
- Analyzes the tokenized input to generate embeddings representing the text's sentiment.

Sentiment Score

- Converts the embeddings into a sentiment score.
- Typically, the sentiment score can be positive, negative, or neutral.

TECH - STACK

- **Programming Language: Python**
- **Pre-trained Models: BERT**
- **Sentiment Analysis Tools: VADER**
- **NLP Libraries: Hugging Face Transformers, NLTK, spaCy**
- **ML Frameworks: TensorFlow, PyTorch**
- **Data Visualization: Power BI**

CONCLUSION

- The project successfully demonstrated the use of web scraping and BERT for sentiment analysis.
- We achieved efficient extraction and classification of customer reviews.
- Identify trends in the sentiment data.
- Highlight common themes in positive and negative reviews.
- Derive actionable insights from the analysis.



TEAM CONTRIBUTIONS

TEAM MEMBERS:

- **PRASUK JAIN (TEAM LEADER)** –nltk models and ml classification
- **V SANJAY KUMAR** –bert model and data visualization
- **PARAM PATEL** –web scrapping and data cleaning
- **ADITI V** –vanders model and process work flow
- **THACKSHANARAMANA B** –documentation and feature extraction

**Thank
you**