TITLE: Fake news detection for Whatsapp Forwards.

TEAM LEAD NAME (STUDENT):
Rakshita Bingi Sai
TEAM MEMBERS:
Shikhar Pandey
Aryan Karoliwal
Riya Kumari
Safeer Ahmed

Answer all the following questions in 5 to 10 bullet points

PROBLEM/PAIN POINT/OPPORTUNITY IDENTIFIED

Explain the Problem/pain point (or) Opportunity Identified with quantification of problem and some research data

1. In the recent times a lot of fake information has been virally spreading through social media platforms which has created a lot of political tension along with fake information being rapidly spread and has resulted in a lot of chaos around the globe.
2. This has led to a lot of conflict among people regarding beliefs and trust including action.
3. People being vulnerable tend to believe the fake information without cross checking it with actual sources and further go on to spread the information
4. Now this problem can be tackled if there is a system to identify or regulate information that is being passed on by untrusted sources
5. During the covid peak , a lot of fake information was being flooded on social media and people were not able to differentiate if it was fake or not. Which caused a lot of unrelated data and information flow.

BRIEFLY DESCRIBE THE SOLUTION/ INNOVATION TO ADDRESS THE PROBLEM/OPPORTUNITY IDENTIFIED

1. The solution to tackle this problem can be building a machine learning model that filters through information and data from actual sources.

2. Recently Twitter has added this feature where the app can cross check information and tell if its fake or reliable.

3. The model will help detect and analyze the fake information and tell if its reliable or not.

4. This way the fake information being passed along with people spreading it will be in control.

5. The model will help avoid conflicts and disputes along communities, people and nations.

HIGHLIGHT THE UNIQUENESS/INNOVATIVE COMPONENT OF THE PROPOSED INNOVATION/SOLUTION:

1. The dataset used for model-training is diverse and indegnious.

2. The training model is been connected to the whatsapp-api using twilio, which gives easy accessibility.

3. The diverse dataset helps in higher accuracy due to no region bias.

4. Allows room for improvement of the trained model further by the implementation of LSTM based neural networks or BERT Models to achieve higher accuracy

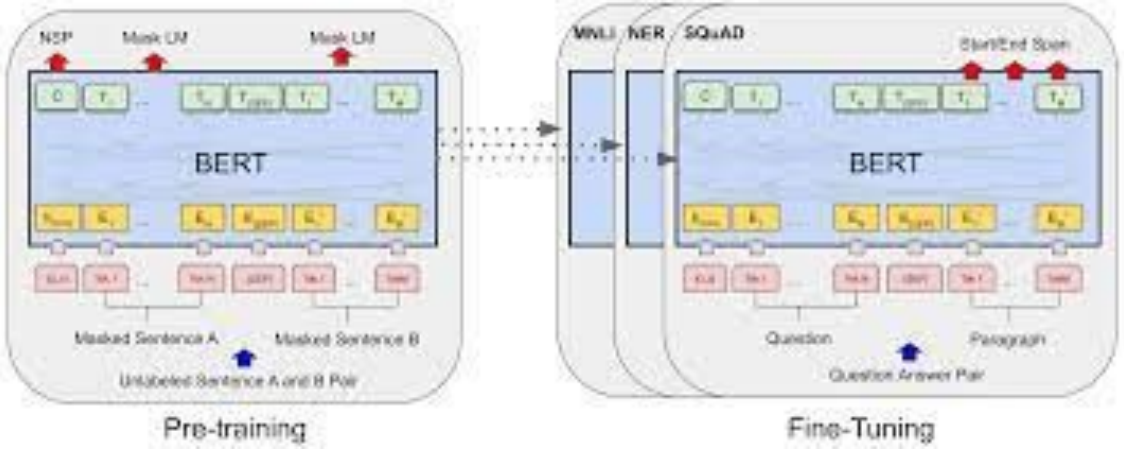
5. The model is automated hence there is no human bias

DESCRIBE THE TECHNOLOGY INVOLVES AND TECHNICAL FEASIBILITY FOR THE PROPOSED INNOVATION/SOLUTION

- Pandas - for data manipulation and dataframe
- Scikit-learn - TfidfVectorizer
- Passive Aggressive Classifier : A machine learning algorithm to train by feefing it instances sequentially in mini batches
- Pickle - to export the trained model

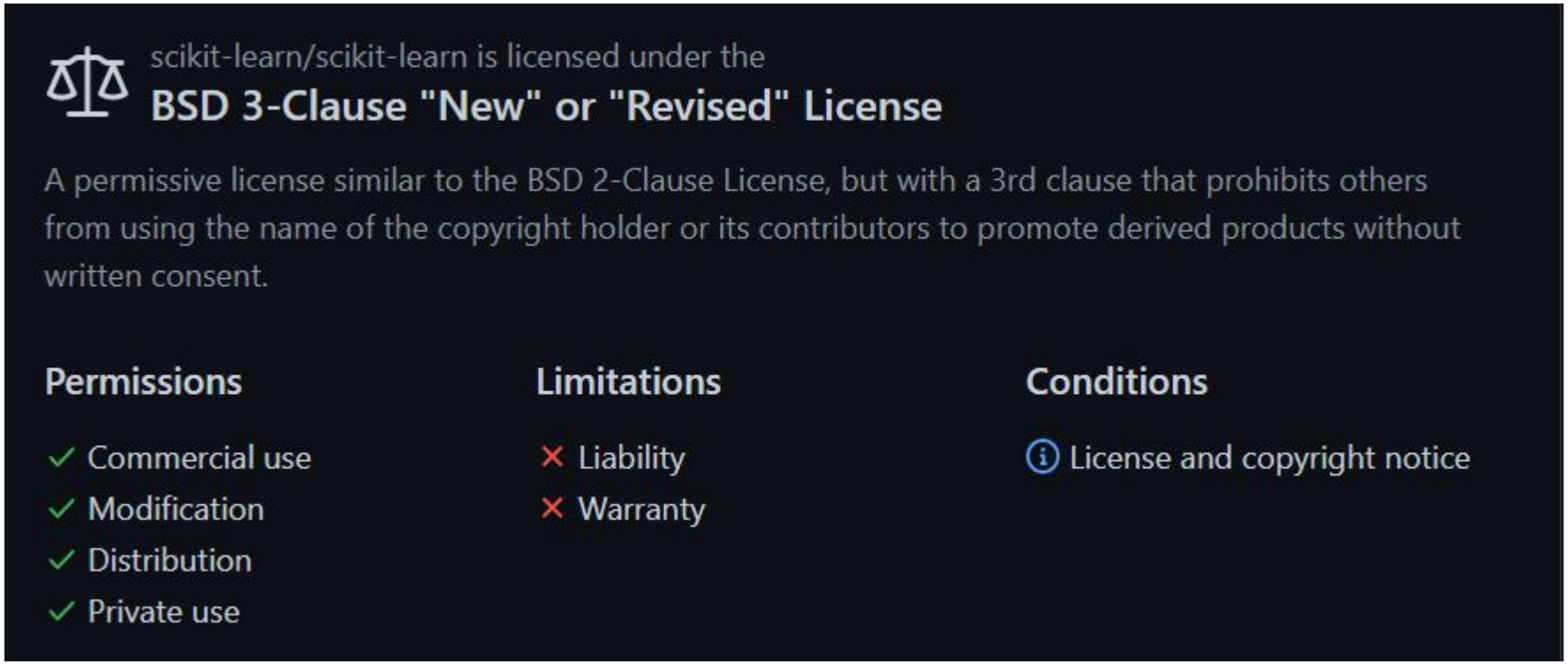
Technological Feasibility

- We look forward to implement the same model in BERT Model basically BERT stands for Bidirectional Encoder Representations from Transformers
- BERT's key technical innovation is applying the bidirectional training of Transformer, a popular attention model, to language modelling.
- This is in contrast to previous efforts which looked at a text sequence either from left to right or combined left-to-right and right-to-left training.



HIGHLIGHT IF ANY INTELLECTUAL PROPERTY (IP) COMPONENT ASSOCIATED WITH THE PROPOSED INNOVATION/SOLUTION.

Scikit-learn is licensed under the BSD 3-Clause



DESCRIBE THE BUSINESS POTENTIAL OF THE PROPOSED INNOVATION/SOLUTION INTO VENTURE/START-UP

- The various platforms that can be used to disseminate content effectively and widely;
- The types of data news article may contain, and the impact of each type of data on readers
- The different types of fake news categories,Existing fake news detection methods,and
- Current data sets that are available for fake news detection.
- Big companies,especially social media platforms would want a fake news detector to get rid of fake news circulating on platforms to get rid of the negativity surrounding their brand names. Also,some fake news sites are clickbaits that steal user's data ,jeopardising the user's security on such platforms.
- This can be implemented as a service across WhatsApp globally to provide a more reliable social media platform.
- A browser plug-in can help identify and hence tag the news as “verified” or “not verified.”

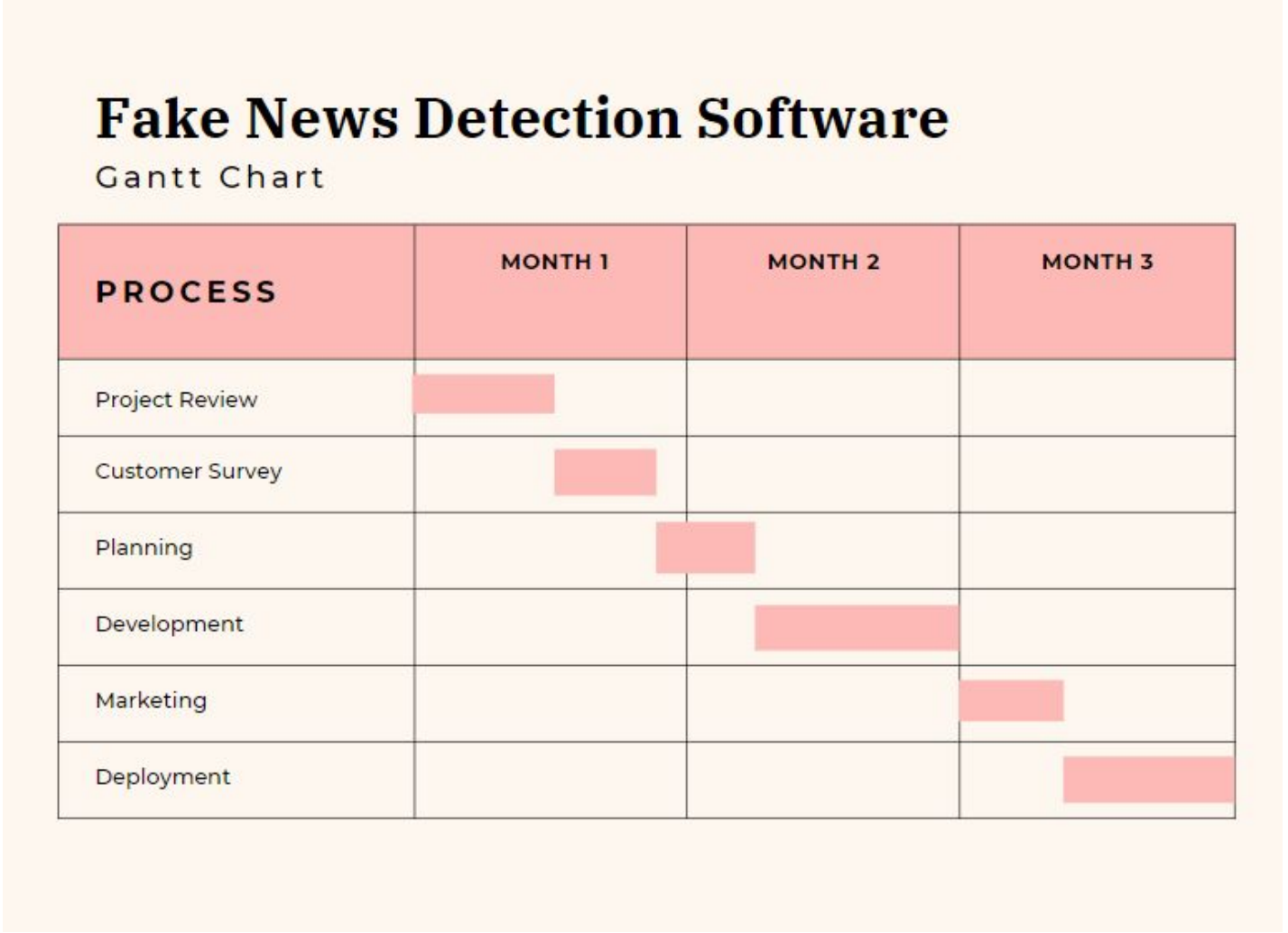
HIGHLIGHT THE MARKET POTENTIAL FOR INNOVATION/SOLUTION

- Most of the market spending goes towards the detection and monitoring part of the Information Warfare, while a smaller portion is invested towards digital forensics, trying to trace the source, after the damage has been done.
- The proliferation of technology, specifically Artificial Intelligence, allow wide scale and easy to use solutions to create synthetic media that seems very real, either in video, photos, or text. In combination with networks of fake profiles, bot networks and smart utilization of social media, the attackers have many ways to influence either the public or specific individuals.
- As media and technology are now being used for cyber terrorism, criminal activity, espionage and military warfare, the new age of information warfare threatens democracies worldwide, pushing governments to combat this threat with large investments.

EXPLAIN THE CONSTRAINTS/RISKS ASSOCIATED WITH THE INNOVATION/SOLUTION AND STRATEGY TO TACKLE THESE

- Completeness of Data: There would some data that will be incomplete which would miss some negation key words and lead to erroneous data.
- Timeliness of Data: Data acquired can be outdated.
- Social Media won’t give access to there message or post as part for of their Privacy Policy
- For every data access, the service will have to request the user for permissions.
- Handling acquired data securely would be troublesome for the developer.

IMPLEMENTATION PLAN WITH TIMELINE TO CONVERT THE INNOVATION/SOLUTION TO A VENTURE/STARTUP



TEAM COMPOSITION AND COMPETENCY AND SKILL SET TO TURN THE INNOVATION/SOLUTION INTO START-UP

Name	Skill	Contribution
Shikhar Pandey	Project Management	Overviewing workflow
Aryan Karoliwal	Data Analyst	Capturing and cleaning of data
Riya Kumari	Data Analyst	Pattern recognition
Safeer Ahmed	Machine Learning	Back End Engineering
Bingi Rakshita Sai	Deep Learning	Back End Engineering

PHOTOGRAPHS: TEAM

THANK YOU