**Team name: Pixel Prophets**

Mentor’s Name: Dr Ranjit Panigrahi

Abstract:

Our hackathon project introduces an innovative browser extension designed to detect and combat dark patterns—manipulative user interface strategies—amidst growing concerns about deceptive design practices on digital platforms. Utilizing advanced algorithms and machine learning, our extension provides real-time feedback to users, empowering them to make informed decisions while promoting an ethical online experience.

Key Features:

* Pattern Recognition: Leveraging machine learning models, our extension identifies a diverse range of dark patterns, from misleading wording to aggressive nudges, ensuring comprehensive coverage.
* Accuracy: Used three transformer based open-source models (BERT, RoBERTa and XLNet) for more accurate detection of dark patterns.
* Real-time Feedback: Users receive instant notifications and visual cues when interacting with websites using dark patterns, fostering awareness and enabling conscious decision-making.
* User Reporting Feature: Empowering users to actively contribute, the extension includes a reporting feature. Users can flag websites engaging in perceived shady practices, enriching the extension's database and fostering a collective effort towards a safer online environment.
* Community-driven Database: Continual learning is facilitated through a community-driven database, allowing users to contribute new patterns and updates, enhancing the extension's detection capabilities over time.
* Compatibility and Accessibility: Seamlessly integrated into popular browsers, our extension prioritizes accessibility, ensuring a broad user base can benefit from a safer and more transparent online experience.

Team Leader Name: Gagan Sarmah

Team Member Names:

1. Grenish Rai
2. Anish Acharya
3. Kunsang Moktan

**Problem Statement**

Design and development of software-based solutions that can detect the use, type, and scale of dark patterns on e-commerce platforms.

**Solution Strategy**

In response to the aforementioned issue, we've introduced a browser extension named **“Pixet”** designed to identify both the types and quantity of dark patterns within a webpage. Employing **transformer-based** open-source language models, including **BERT**, **RoBERTa**, and **XLNet**, Pixet is capable of detecting and highlighting the presence of dark patterns on any Chromium-based browser.

**Tools and Methods Used**

* Extension: HTML, CSS, and Vanilla JavaScript
* Models: Python (Transformers, Pytorch, Pandas other python libraries)
* Content Scraping: Node.js (Puppeteer, Express, Cors, Child\_Process)
* Tech Stack: Jupyter Notebook, GitHub, VS code

**Result and Discussion**

The problem at hand has been addressed through the development of our browser extension, **“Pixet”** This extension effectively identifies and quantifies various dark patterns present on webpages. By leveraging transformer-based open-source language models like **BERT**, **RoBERTa**, and **XLNet**, Pixet seamlessly operates across all Chromium-based browsers.

As of now, Pixet excels in detecting six text-based dark pattern categories: Scarcity, Misdirection, Urgency, Social Proof, Forced Action, and Sneaking. Our ongoing efforts include the expansion of our dataset and the integration of additional dark pattern categories, enhancing the extension's ability to detect and combat deceptive practices on the web.

**Future Scope**

1. Deployment of Extension
2. Collection of more datasets to train the models
3. Making the model learn by itself using user’s feedback in real time
4. Using various technologies like OCR (to extract text from images) and site tracking for better detection of dark patterns

**Acknowledgement**

*This project is based on the Mathur et al.’s study and its dataset. We thank their authors for making the source code publicly available.*

*We have also used the dataset from YadaYuki's workshop on IEEE BigData 2022 which was also updated from Mathur et al's dataset.*