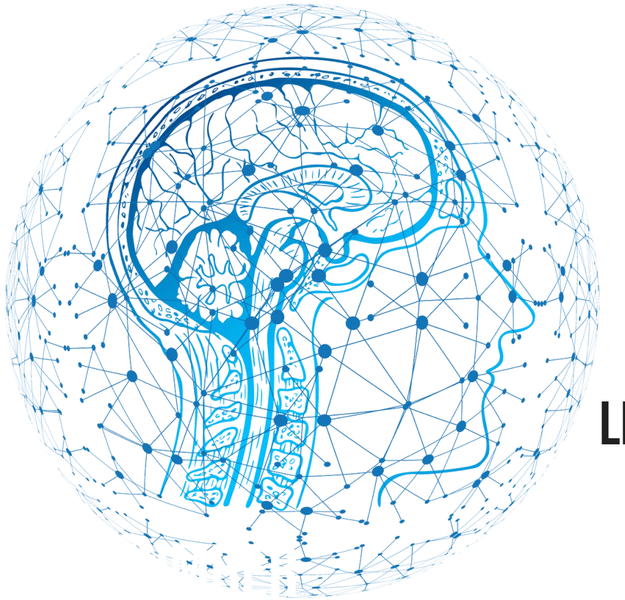
Ethical AI Recommender Systems For IPhones

Hands-on Project



LEVERAGING AGENTIC AI, PROMPT ENGINEERING, ETHICAL TOOLS, RAG, LLM, ML AND NLP

Team Number: AAI-531 Group 7

Team Members :

Gangadhar Singh Shiva,

Ananya Chandraker,

Harish Kapettu Acharya

Project Objectives - AI Ethics Principles

- Build a **transparent, fair**, and privacy-preserving recommender system
- Ensure that recommendation outcomes are **interpretable, unbiased, and protect user privacy**
  - Address **bias and explain model** predictions using **SHAP and LIME**
  - Use explainability tools to detect and mitigate potential sources of discrimination or **unfairness in model behavior**

Accountability: Clearly defining responsibility for decisions and outcomes of the recommender system.

- Utilize modern AI tools including RAG, transformers, and differential privacy
- Leverage state-of-the-art methods to enhance recommendation quality while **embedding ethical AI safeguards**

Unintended Stakeholders and Impacts

Potential unintended stakeholders: Marginalized or minority demographics, competitors, third-party developers, and international users.

Impacts:

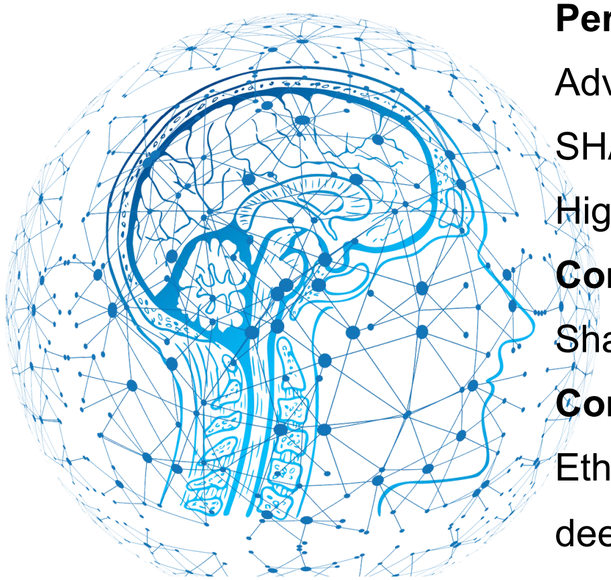
Risk of unintentionally reinforcing existing biases leading to discriminatory or unfair recommendations.

Potential violation of user privacy through insufficiently secured data practices.

Unfair competitive dynamics resulting from biased or non-transparent algorithms.

Ethical Harms: Potential discrimination, compromised user privacy, loss of consumer trust, and reduced market fairness.

ETHICAL EXPLAINABILITY AND TRANSPARENCY



Personal Perspective:

Advocates strongly for transparent and interpretable recommendations using SHAP and LIME AI Ethic Tools.

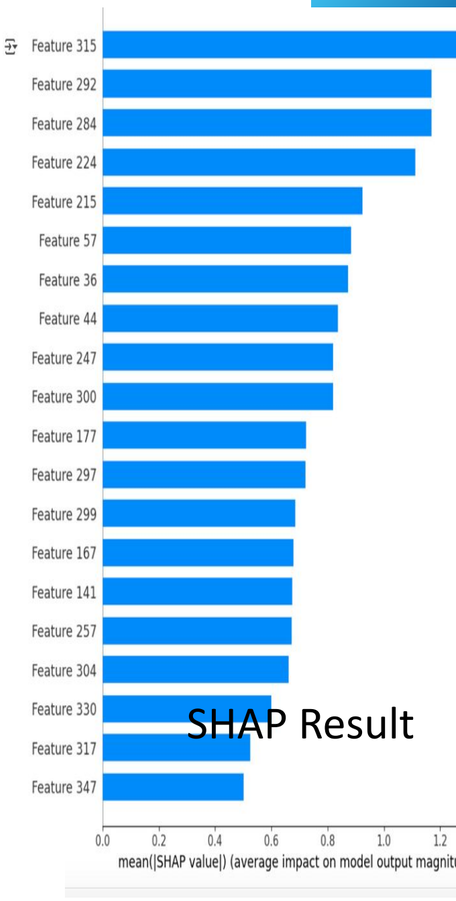
Highlights fairness as a cornerstone to ensure unbiased outcomes.

Comparison to Team:

Shares transparency focus, with unique emphasis on interpretability tools.

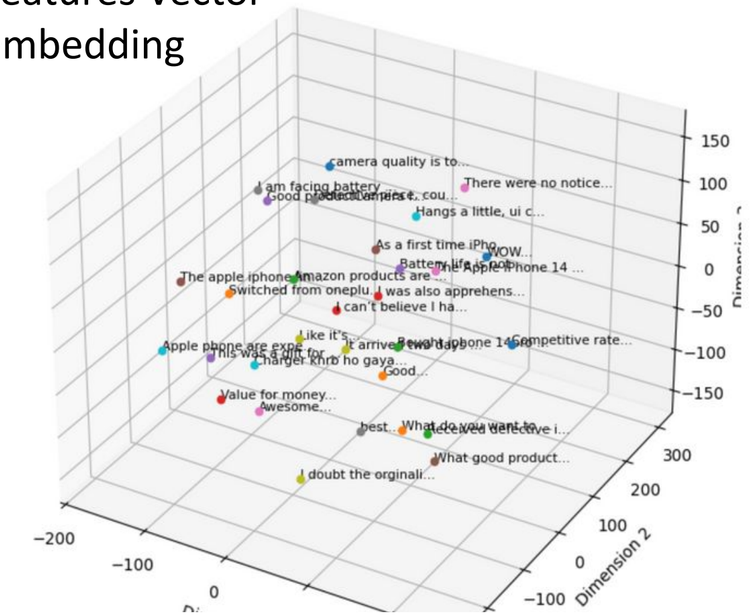
Common Sense View:

Ethical recommender systems must prioritize transparency and fairness, rooted deeply in personal values of integrity and equality.

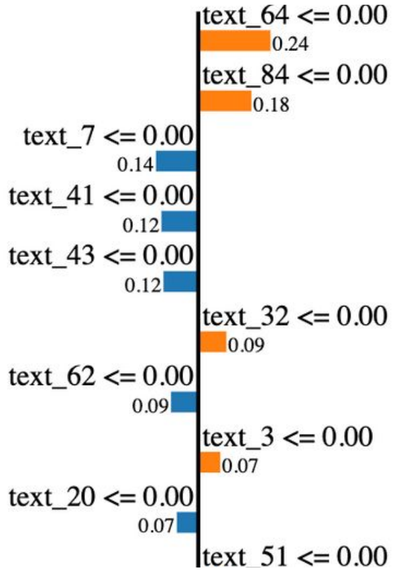


SHAP Result

Features Vector embedding

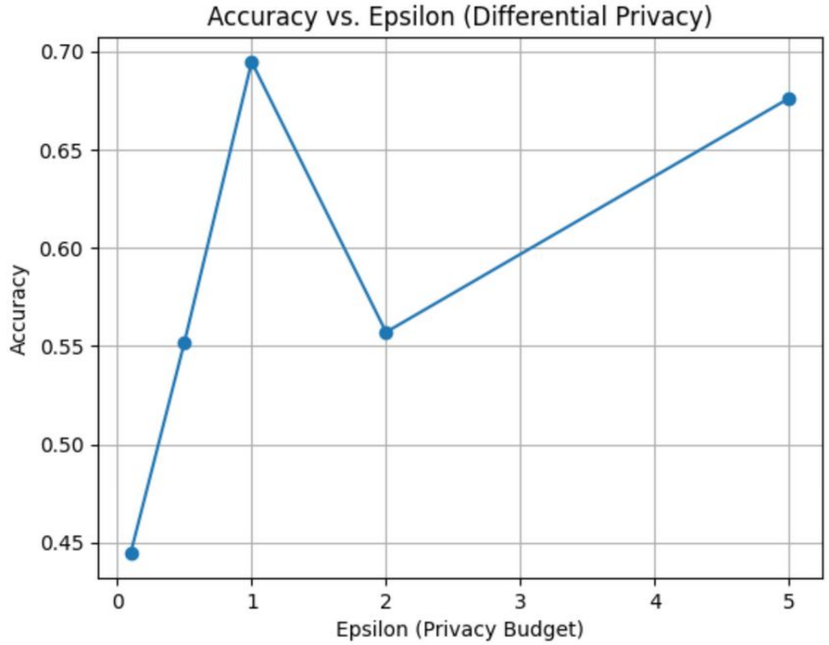


LIME Result



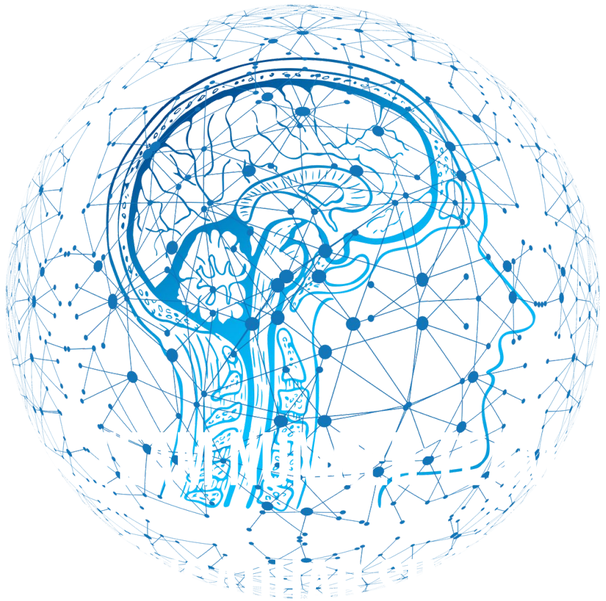
ETHICAL - PRIVACY ENHANCEMENTS

- **Personal Perspective:**
- Emphasizes fairness techniques, such as demographic parity, to prevent biased outcomes.
- Advocates strongly for practical ethical implementations.
- **Relation to Team:**
- Strong agreement on fairness and privacy, with practical focus on ethical accountability.
- **Common Sense View:**
- Ethical AI recommendations must actively promote fairness and unbiased treatment, aligning with values of equality and ethical responsibility.




```
Accuracy with varying Epsilon (Differential Privacy):
 - Epsilon: 0.1, Accuracy: 0.44
 - Epsilon: 0.5, Accuracy: 0.55
 - Epsilon: 1.0, Accuracy: 0.69
```

ETHICAL PRINCIPLES (FAIRNESS)



Personal Perspective:

Strong emphasis on data privacy, utilizing differential privacy and thorough data sanitization practices.

Views privacy protection as essential to preventing ethical harms.

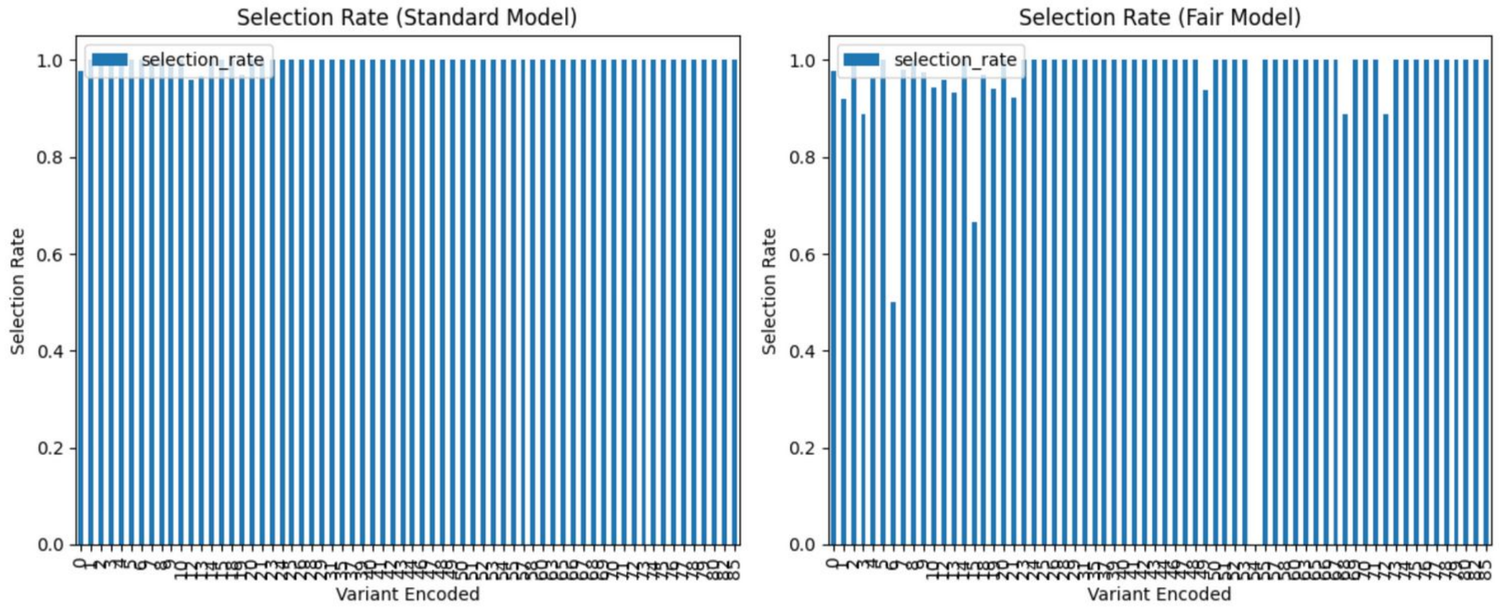
Contrast to Team:

Agrees broadly on transparency but emphasizes rigorous privacy measures.

Common Sense View:

Protecting user privacy is paramount, aligning with personal values of autonomy, respect, and accountability.

```
Comparison:
- Standard Model DP Difference: 0.0408
- Fair Model DP Difference: 1.0000
```



CONCLUSION

- Reinforced importance of SHAP and LIME for transparency and interpretability.
-
- Emphasized fairness and differential privacy as critical elements ensuring equitable and secure AI recommendations.
-
- Team consensus highlights the necessity of fairness, transparency, privacy, and accountability.
-
- Advocates continuous ethical review, responsible governance, user empowerment, and adherence to regulatory frameworks for maintaining user trust and ethical integrity.

THANK YOU