DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION SYSTEMS

INTRODUCTION TO AI

FINAL PROJECT : WRITTEN REPORT ONE

THE INK PULSE PROJECT

MEMBERS

AGATHA SUZIE SIEYAPNJI YOUYOU(91292027)
BEN RIAK CHANGDAR(99392027)
FAVOUR AMOURZANG FRI FON(31202027)
JOSEPHINE FORGIVE DOAMEKPOR(41892027)

THURSDAY, IST MAY, 2025

Background: Introduction to the Organization and its Problem

Ashesi University makes commendable efforts to attentively listen to the needs of the student body and actively works to meet those needs. Over the years, various support units such as the Office of Diversity and International Programs(ODIP), the Counselling, Coaching, Advising and Peer Support (CCAPS), have been established to help students thrive academically, professionally and personally. These help units, often in collaboration with the Ashesi Student Council(ASC) curate events, workshops and initiatives aimed at supporting students and enriching their campus experience. They also consistently seek feedback to ensure that their services remain relevant and impactful.

However, despite these efforts, student engagement with these resources remains surprisingly low. Event turnouts are very low, leading to significant time and resource wastage. Our hypothesis is that this lack of engagement is due to the creation of events that do not exactly meet the pain points of students.

Evidence supporting this hypothesis comes from anonymous submissions to The INK, a student- run Instagram platform that allows students to share unfiltered reflections about life at Ashesi. Submissions, collected via a Google form, often touch on critical issues such as academic pressure, mental health struggles, and general student life experiences.

To further validate our observations ,we spoke with the head of the CCAPS department and representatives of ODIP who confirmed that many events suffer from low attendance and that resources are often underutilised. Both departments acknowledged the challenge in designing support services that students feel are genuinely relevant and accessible. .

Applying Machine Learning to solve the Problem

Our problem Statement: Although Ashesi University has made commendable efforts to support its student body through various help units, there remains a persistent gap between the services provided and the actual, pressing needs of students which is evidenced by low engagement levels and poor event attendance. This is largely due to the lack of structured data driven insights into what students are truly experiencing and struggling with. Without a systematic way of identifying these concerns, support units are left to rely on limited feedback mechanisms which often miss deeper and recurring themes in student life.

To bridge this gap, we are leveraging Machine learning, specifically Topic Modelling, (An unsupervised machine learning technique that can automatically uncover hidden themes in large volumes of text data),to analyse the anonymous student submissions shared on The INK. By applying algorithms such as Latent Dirichlet Allocation (LDA) and BERTopic, we can identify common concerns, recurring pain points and emerging trends without requiring anyone to categorize these experiences themselves.

This approach offers several key benefits including:

- *Improving targeted support design*: Help units can then use these insights to design events, outreaches , etc that directly address students' most pressing concerns and will subsequently have high attendance.

- *Scalable Analysis*: Hundreds of submissions can be analyzed quickly and efficiently , revealing latent patterns that may go unnoticed through manual reading.

- *Sustainability*: As new submissions come in, topic modelling can be rerun to track the shifts in students' concerns over time..

Data Collection and Preprocessing Challenges
Our initial plan was to use Instaloader, a python tool, to extract public posts from The INK Instagram page, along with their engagement data(likes and comments). This would have added an extra layer of insight by showing which themes resonated most with students. However, due to Instagram's updated privacy restrictions, this method was blocked and required account authentication, raising ethical concerns.

As an alternative and more ethical route, we reached out directly to the student who manages The INK and explained our project. Thankfully, he agreed to share a spreadsheet of anonymized student submissions, including timestamps and free- text responses from 2022.

Our biggest challenge with data processing was the informal nature of some student submissions. Many responses were written in slang or contained pidgin english phrases, which posed difficulties for accurate topic modelling. To address this, we built a Pidgin-to-English converter using the Gemini API, which helped translate informal and non standard expressions to more formal english, ensuring that our data was cleaner, consistent and ultimately improved the quality of our results.

However, we faced an issue with working within Gemini API's rate limits (*Limites De Débit*, n.d.). Since we were using the free plan, the service enforces strict request limits per minute and per second. This made translating our entire dataset particularly difficult, as the API would throttle requests once these quotas were exceeded. To work around this limitation, we implemented two key solutions: first, we broke the data into smaller chunks to stay under the maximum request size limits, and second, we added strategic time delays between API calls (2-second intervals) to comply with the rate restrictions. While these adjustments allowed the translation to complete successfully, they extended the total processing time to approximately 10 minutes for our full dataset.

Ethical Considerations

Given that the data involved anonymous student opinions, ethical responsibility was a key consideration throughout this project.

- Data Privacy: Although the responses were anonymous, we prioritized reaching out to the page's administrator and obtaining permission before using the data. We ensured

that no identifying information was included in the analysis or visualizations. Our goal was not to expose individuals but rather amplify collective student responses in a responsible and respectful manner.

● Informed Consent: While students voluntarily submitted feedback, it is important that The Ink clearly communicates how responses may be used for analysis in future forms.

● Use of Results: Our goal is not to critique individuals or departments but to surface insights that can guide constructive improvements and for that reason we are intentional about how results are used and interpreted by ensuring a focus on themes only.

Limites de débit. (n.d.). Google AI for Developers. https://ai.google.dev/gemini-api/docs/rate-limits?hl=fr