Team Insomniac

- Shreya Singh | 2022 | Birla Institute of Technology - Mesra, Ranchi

- Aastha Singh | 2022 | Birla Institute of Technology - Mesra, Ranchi

- Tanweerul Haque | 2022 | Birla Institute of Technology - Mesra, Ranchi

- Akshat Dubey | 2023 | Birla Institute of Technology - Mesra, Ranchi

Tell us a bit about yourself

- **Any projects you've worked on -**
  1. **Saving Newborns Life through Artificial Intelligence** - Demonstrated via a retrospective study that the infant cry provides rich source of information about the health of a newborn. Developing a prototype mobile app which could be used in resource-poor settings using techniques from automatic speech recognition, Machine Learning and Deep Learning. Mel frequency cepstral coefficients (MFCC) feature extraction technique was used to extract features from baby's cry.
  2. **Early Sepsis Detector** - Implemented ensemble based classifier models on a highly imbalanced physiological dataset. Trained model was deployed as a Rest API using Flask on Heroku to be consumed by an Android frontend.

- **Past Hackathon experiences -**
  1. Mentors - Hack the Mountains, Semester Long Projects Mentor - DAIICT-Gandhinagar.
  2. President (Society for Data Science, Team Aveon Racing - BIT Mesra) and Vice President (Society for Data Science, BIT-Mesra)

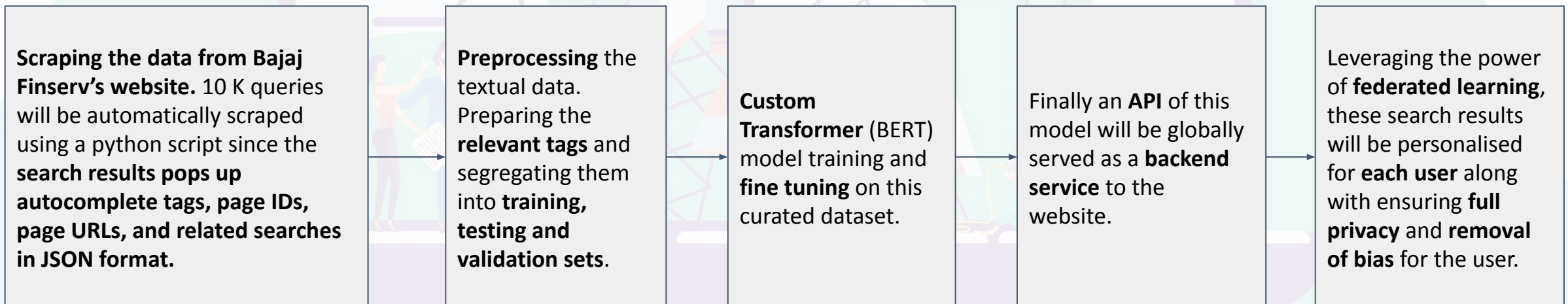- **Accolades or awards that you have received -**
  1. Saving Newborns Life through Artificial Intelligence stood first position in the nomination among top three qualified teams for the finals of National Innovation Competition, Ministry of HRD.
  2. Early Sepsis Detector was part of Smart India Hackathon (SIH)-2020 and was recommended for finals.
 3. Google Summer of Code (GSoC) - 2021

RECOMMENDATION ENGINE

We will first scrap the data from the website. Then, we will create a dataset containing two columns named user-query and the respective keywords which we obtained after scraping. As the next step, we will fine-tune a transformer based model (BERT: Bi-Directional Encoder Representation from Transformers). After, training of the model we will serve it as an API which will take user-query as an input and will produce most appropriate keywords in the form of JSON file.

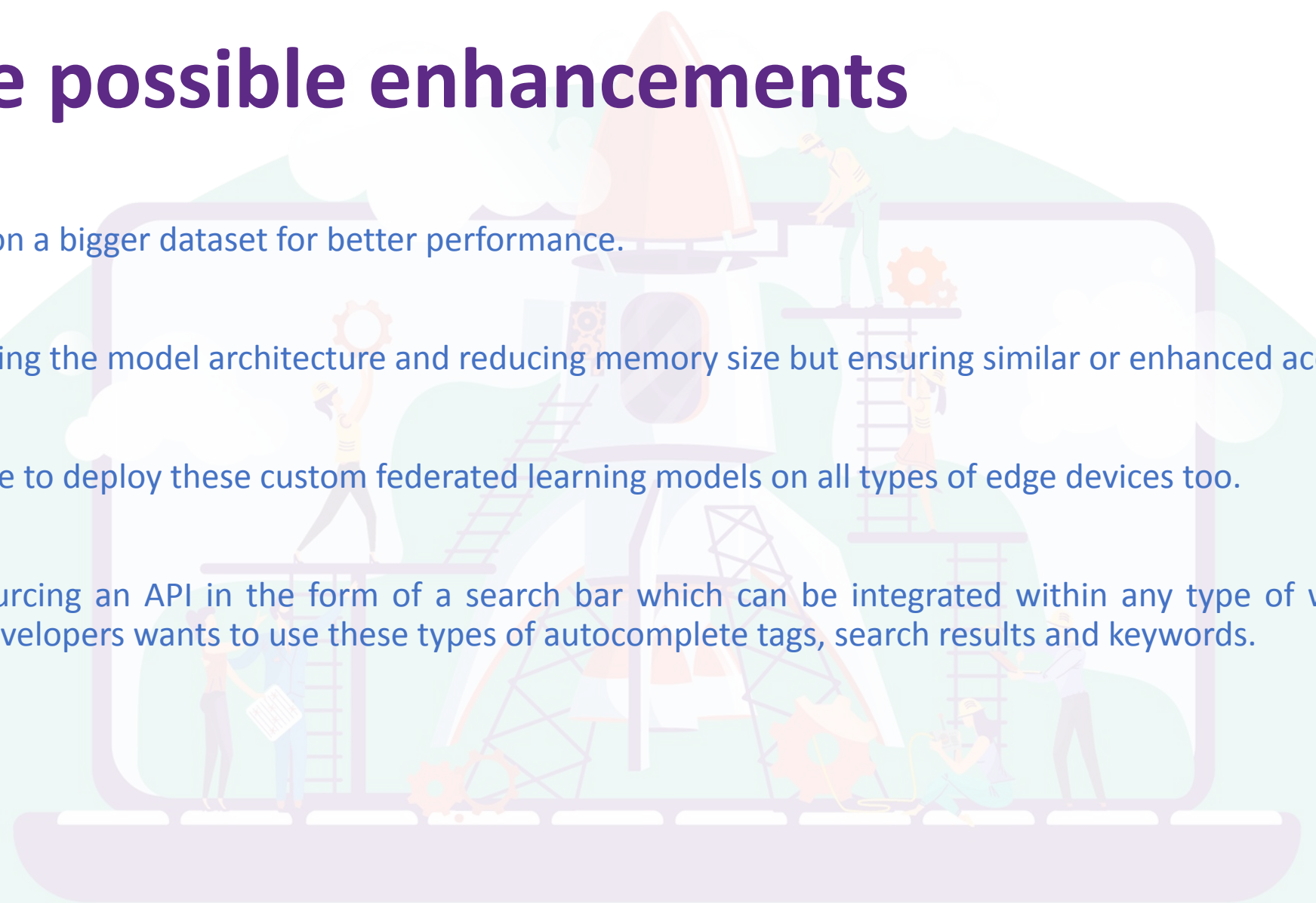|				
Scraping the data from Bajaj Finserv's website. 10 K queries will be automatically scraped using a python script since the **search results pops up autocomplete tags, page IDs, page URLs, and related searches in JSON format.**	**Preprocessing** the textual data. Preparing the **relevant tags** and segregating them into **training, testing and validation sets**.	**Custom Transformer** (BERT) model training and **fine tuning** on this curated dataset.	Finally an **API** of this model will be globally served as a **backend service** to the website.	Leveraging the power of **federated learning**, these search results will be personalised for **each user** along with ensuring **full privacy** and **removal of bias** for the user.

Tech Stack

- **Backend** : Flask, Tensorflow, Python, Javascript, Django

- **Frontend** : Flutter, Python, Django, HTML, CSS

- **Database** : RethinkDB

- **Other Details** : Git, DVC (Data Version Control), Docker
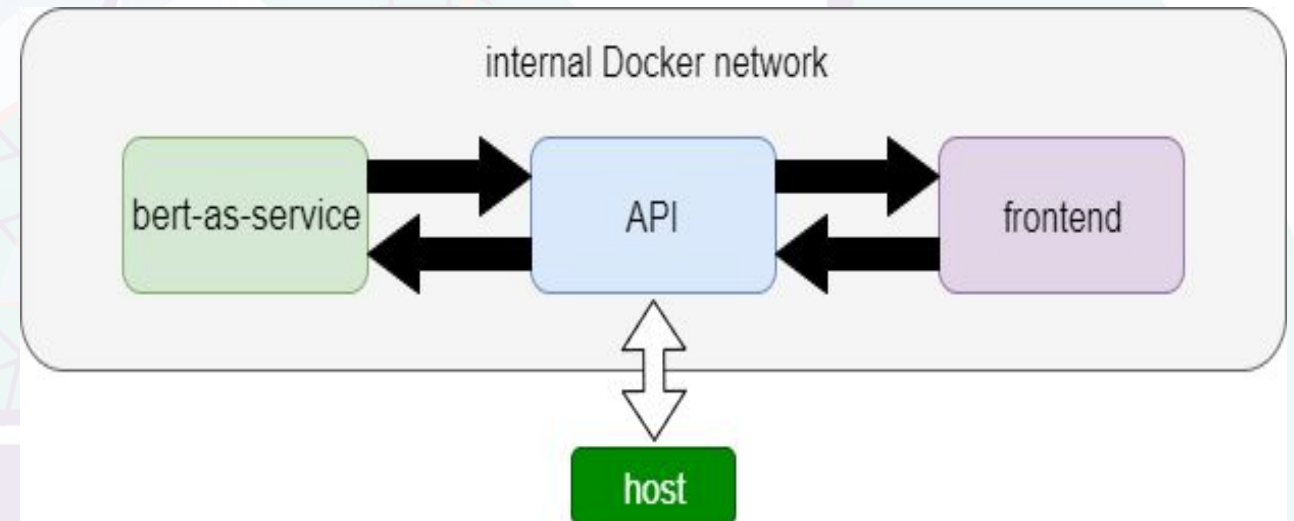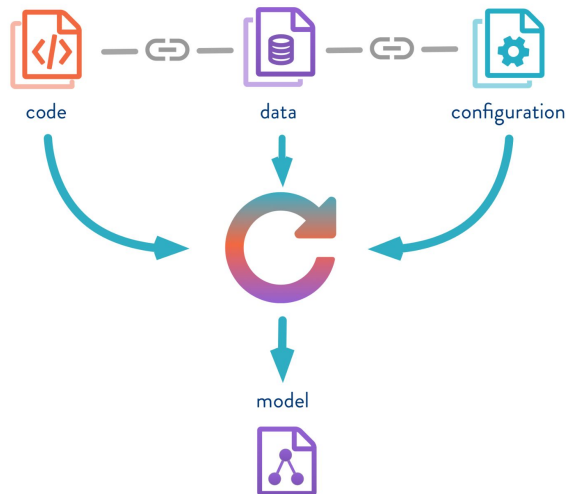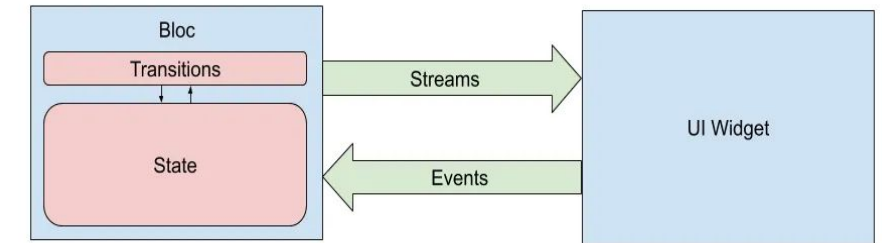
So, how is your solution different?

1. Because of the presence of attention layers, we will utilise a **transformer-based design with attention layers** which enable us to create user inquiries independent of how various users phrase their questions. Because the search results will be fully dependent on **user-specific queries** as well as **global inquiries** as a whole, federated learning will eliminate any form of bias.

2. **Privacy:** Federated learning ensures **full privacy of the user** since there is no need to send the data to a centralized server.

3. **Personalized results** pops up for each user.

4. Our Model will be served as an **API** which can be easily accessed and integrated within the existing website.

5. Our model doesn't use traditional NLP methods/models. Instead, we use a modern and more accurate transformer architectures because of the presence of attention layers. **Model will continue to learn and the performance always gets better over time.**

Future possible enhancements

1) Training on a bigger dataset for better performance.

2) Customizing the model architecture and reducing memory size but ensuring similar or enhanced accuracy.

3) Being able to deploy these custom federated learning models on all types of edge devices too.

4) Open Sourcing an API in the form of a search bar which can be integrated within any type of website where developers wants to use these types of autocomplete tags, search results and keywords.

Risks/ Challenges / Dependencies

- Please mention any risks or challenges that you foresee :

  1. Sometimes there may be irregular response time due to certain reasons.
  2. Re-training of the entire model from scratch will take long time.
  3. Data acquiring is a very hectic process and we may need to assume a particular hypothesis related to the dataset and the features involved.

- Showstoppers in our project :

  1. **Federated Learning**: Federated learning is a machine learning method that enables machine learning models obtain experience from different data sets located in different sites (e.g. local data centers, a central server) without sharing training data. This allows personal data to remain in local sites, reducing possibility of personal data breaches.
  2. **Deep Learning:** Deep learning is an artificial intelligence (AI) function that mimics the human brain's processing of data and pattern creation in order to make decisions.
  3. **Transformer Architecture based model (BERT):** BERT will continue revolutionizing the field of NLP because it provides an opportunity for high performance on small datasets for a large range of tasks. Transformer is a famous attention mechanism that uses artificial intelligence to understand contextual relationships between words in a text.
  4. **The entire model is served as an API.**

Anything Else ?

1. The reason we are emphasizing on using **Federated Learning** in our project as it already offers significant advantages over traditional, centralised approaches, such as keeping the training dataset on the devices, **eliminating the need for a data pool**, and facilitating access to **heterogeneous data** even when data sources can only communicate at certain times. There is no need to aggregate data for continuous learning because models are constantly upgraded using client data. Because federated learning models do not require a **single complex central server to interpret data**, this approach requires less complex hardware, thereby, ensuring **full privacy and removal of bias** from the user queries.

2. Also we are planning to use BERT as **transformer-based design.** It is the best method in NLP to understand **context-heavy texts** as it uses machine learning to predict missing words in text, and examines every sentence without bias, it does a better job than earlier NLP methodologies, such as embedding methods, at comprehending the meaning of homonyms.