LEAP powered by Intel® oneAPI AI Analytics Toolkit

Problem Statement : **Open Innovation in Education**

Team Name : C5ailabs

Team Members : Rohit Sroch, Sujith R Kumar, Mohan K Rachumallu, Shubham Jain

Problem Statement

MOOCs (Massive Open Online Courses)	200K	380M	34.26%	5% -10%
	Users in 2012	Users in 2020	CAGR 2022-27*	Completion rate

Key Challenges

- Lengthy videos
- Instructor Availability
- Slow response from forums
- No real time Q&A/Mentor

Approach

LEAP

(Learning Enhancement and Assistance Platform)

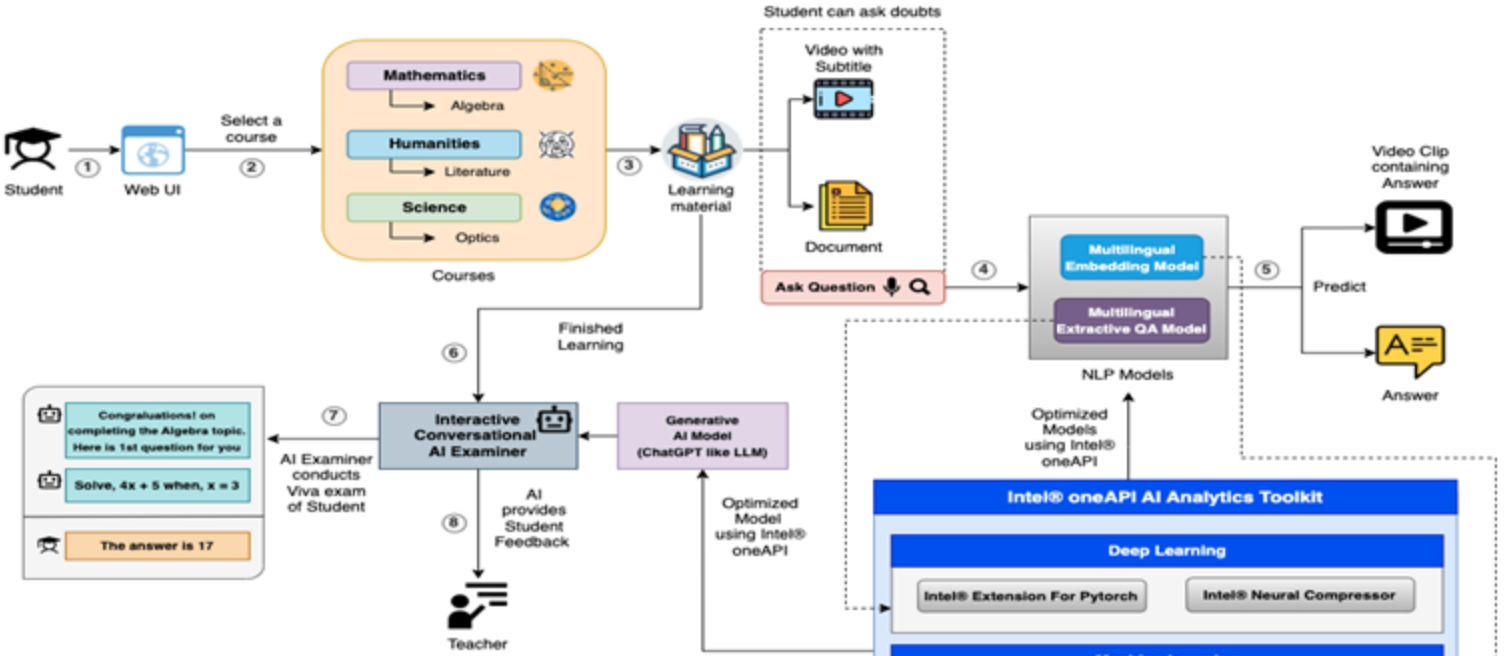
- AI based platform
- Powered by Intel OneAPI
- Quality Education
- All time Availability

Key Features of LEAP

- Ask Question/Doubt
- Conversational AI Examiner
- Feedback from AI Examiner
- Multilingual Support

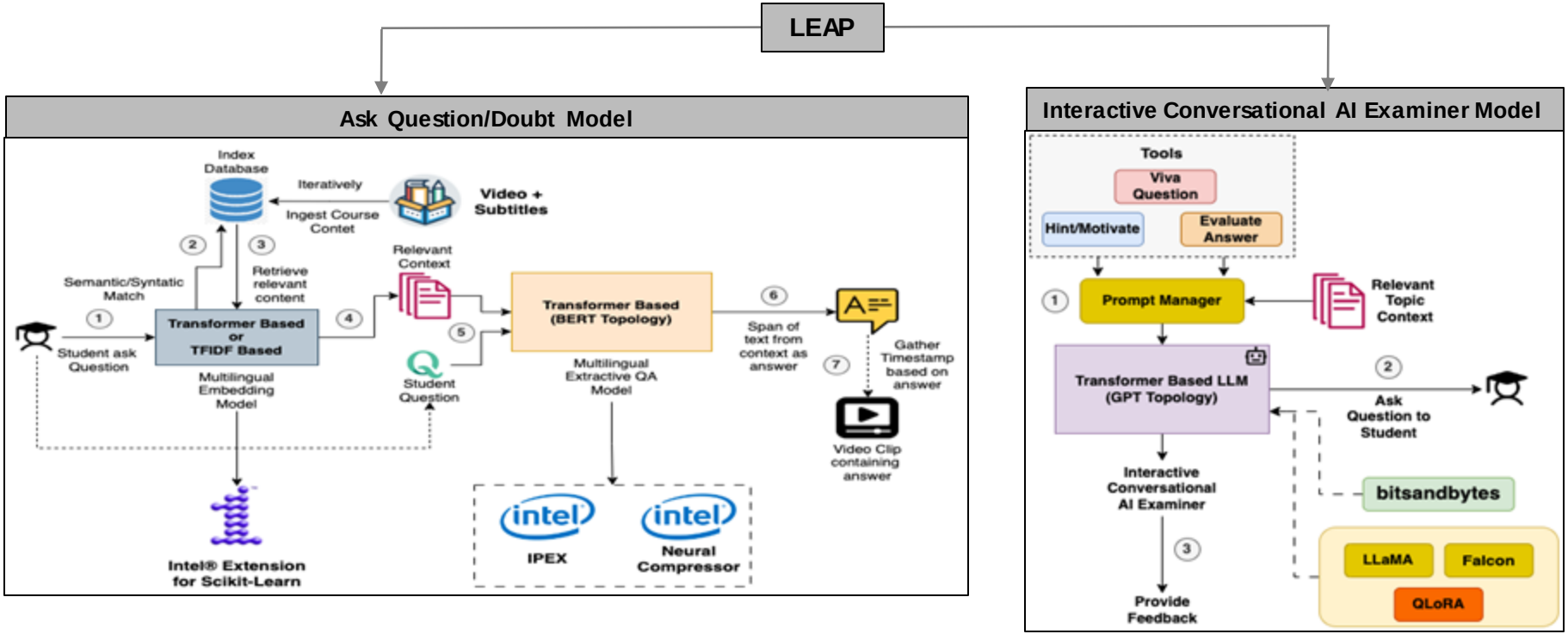
Reference: PRNewswire , Edtechreview; holonig

High Level Architecture

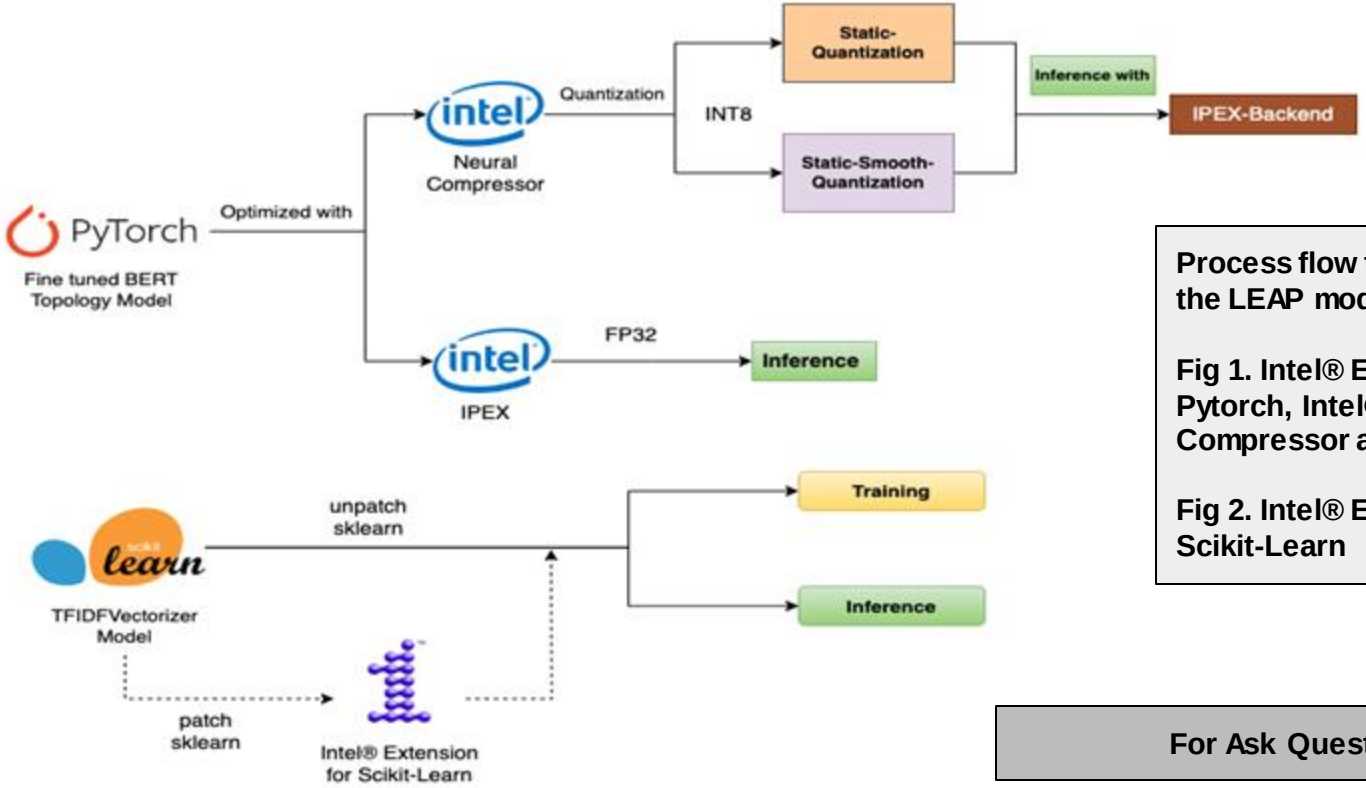


Here Intel® oneAPI AI Analytics Toolkit helped us to optimize the performance of LEAP by faster Inference and Training

LEAP: Detailed Model Architecture Diagram for Both Components



Result Summary (unique aspects of oneAPI/SYCL used)



Process flow to optimize the LEAP models by using

Fig 1. Intel® Extension for Pytorch, Intel® Neural Compressor and

Fig 2. Intel® Extension for Scikit-Learn

For Ask Question/Doubt

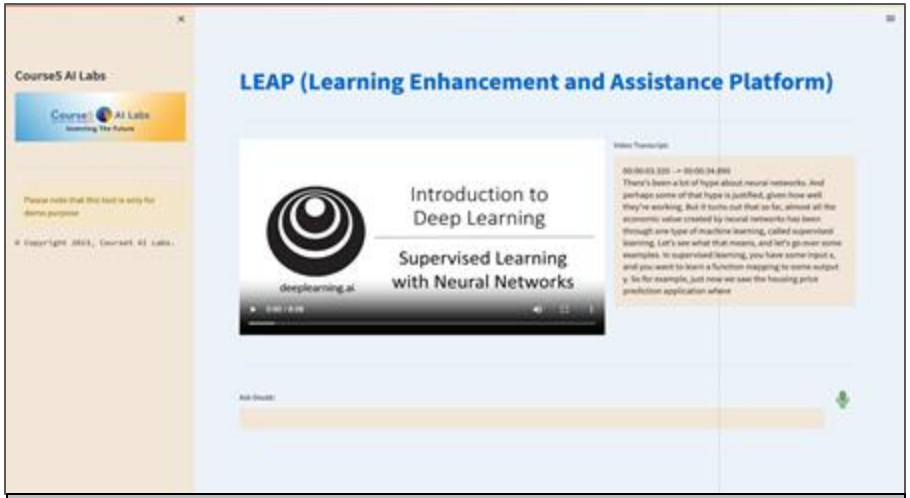
Demo Link and Screenshots

Link: https://www.youtube.com/watch?v=CXkR5tklZm0

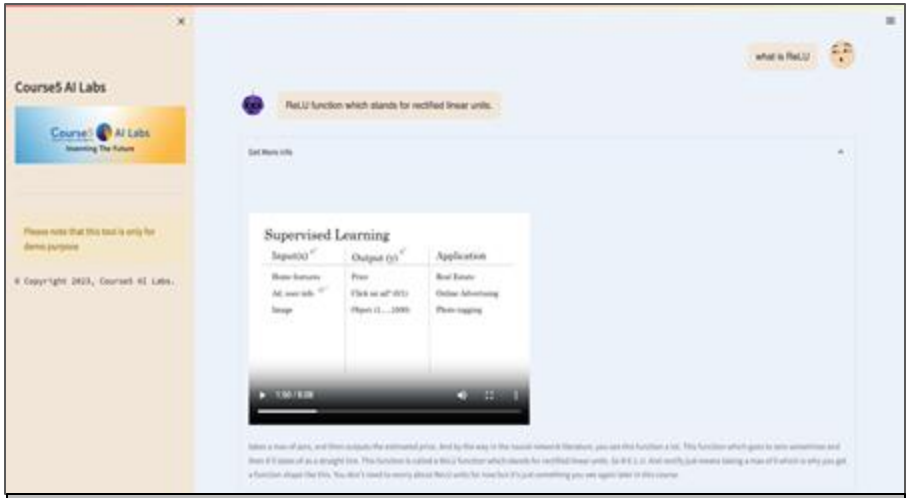
Screen-1 (Dashboard Page)

Screen-2 (Week wise Content Page)

Demo Link and Screenshots



Screen-3 (Ask Question/Doubt Page)



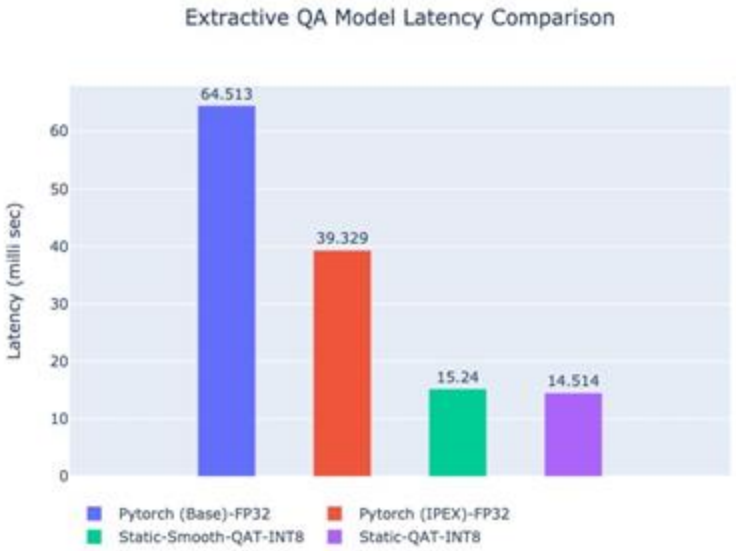
Screen-4 (Ask Question/Doubt Page)

Demo Link and Screenshots

Screen-5 (Interactive Conversational AI Examiner Asks Question to student)

Screen-6 (Interactive Conversational AI Examiner provides hints and motivates a student in case of a wrong answer)

Extractive QA Model (BERT Topology) Latency/Speed-Up Comparison with IPEX and Intel® Neural Compressor



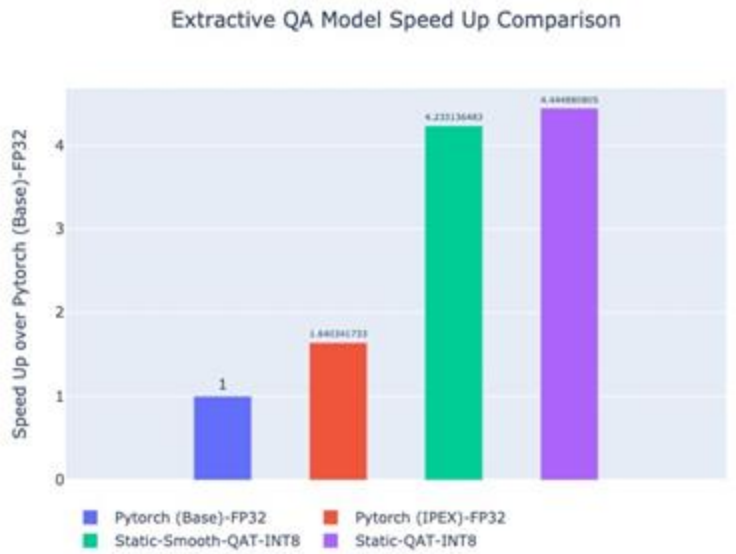
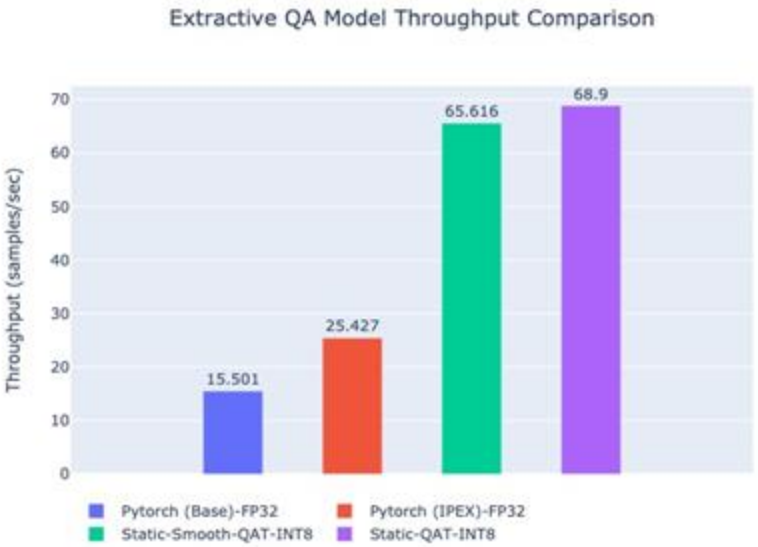


*Fig: Latency/Speed-Up Benchmark result for **our Extractive Question Answering ALBERT Model (Multilingual)** on Intel® Dev Cloud machine (**Intel® Xeon® Platinum 8480+ (4th Gen: Sapphire Rapids) - 224v CPUs 503GB RAM**) with optimization using IPEX-FP32 and Static INT8-Quantization using Intel® Neural Compressor.*

For Ask Question/Doubt Extractive QA Model

Extractive QA Model (BERT Topology) Throughput/F1 Score Comparison with IPEX and Intel® Neural Compressor



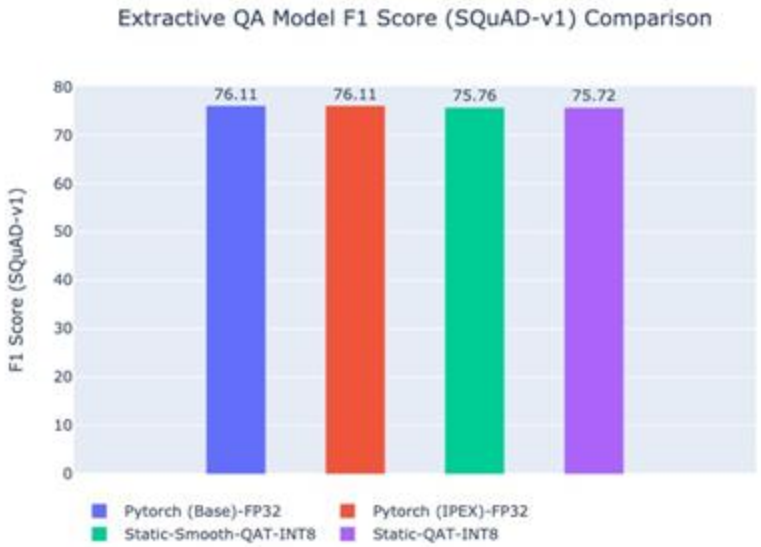


*Fig: Throughput/F1 Score Benchmark result for **our Extractive Question Answering ALBERT Model (Multilingual)** on Intel® Dev Cloud machine (**Intel® Xeon® Platinum 8480+ (4th Gen: Sapphire Rapids) - 224v CPUs 503GB RAM**) with optimization using IPEX-FP32 and Static INT8-Quantization using Intel® Neural Compressor. Also, the model (https://huggingface.co/ai4bharat/indic-bert) was fine-tuned on SQuAD-v1 dataset.*

For Ask Question/Doubt Extractive QA Model

Scikit-Learn (Base) vs Intel® Extension for Scikit-Learn

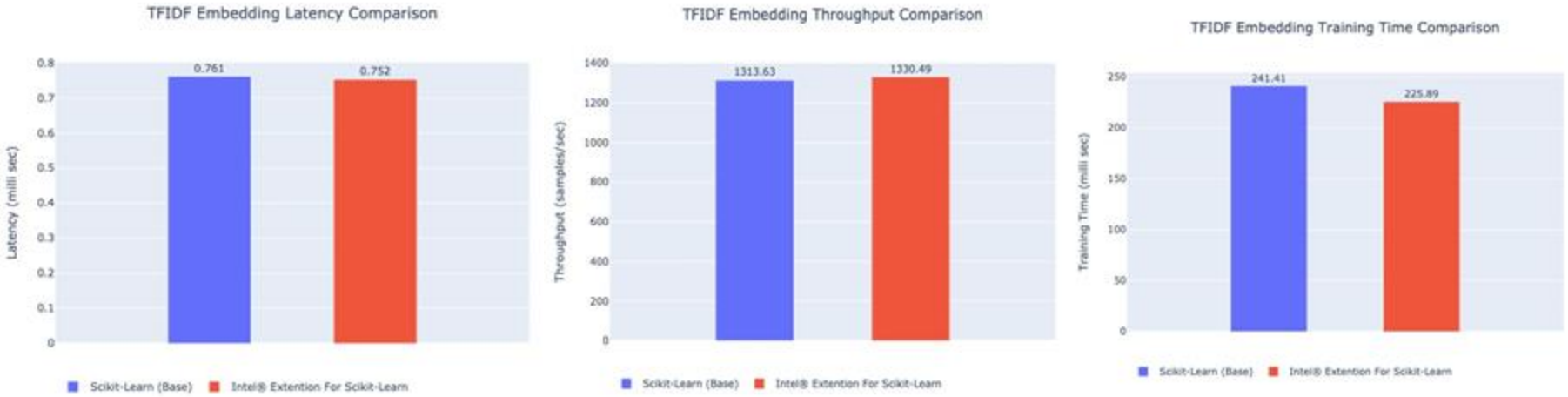
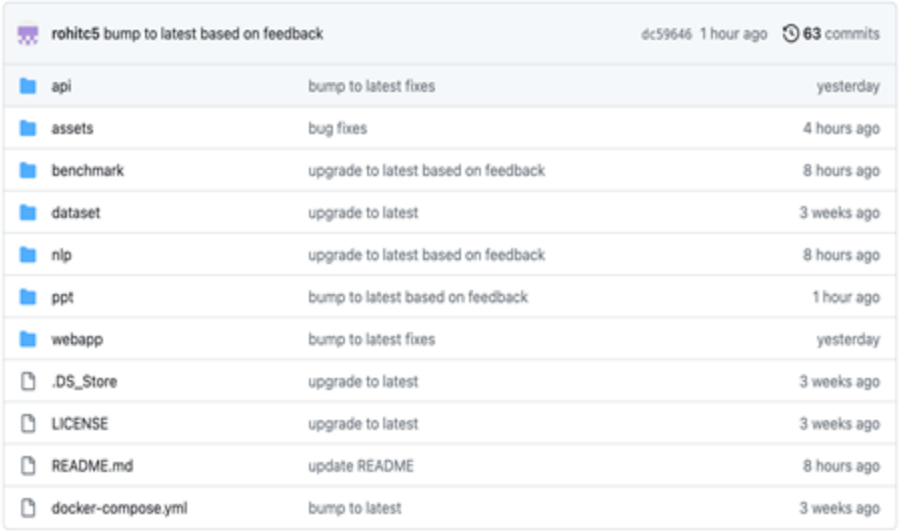


*Fig: Benchmark results for **TFIDFVectorizer** Embedding model during training and inference on Intel® Dev Cloud machine (**Intel® Xeon® Platinum 8480+ (4th Gen: Sapphire Rapids) - 224v CPUs 503GB RAM**). Please Note that we don't see much of a difference may be because we used a tiny dataset.*

For Ask Question/Doubt Embedding Model

GitHub Link (Codes should be public and available after hackathon also)

https://github.com/rohitc5/intel-oneAPI



rohitc5 bump to latest based on feedback — dc59646 1 hour ago — 63 commits

api	bump to latest fixes	yesterday
assets	bug fixes	4 hours ago
benchmark	upgrade to latest based on feedback	8 hours ago
dataset	upgrade to latest	3 weeks ago
nlp	upgrade to latest based on feedback	8 hours ago
ppt	bump to latest based on feedback	1 hour ago
webapp	bump to latest fixes	yesterday
.DS_Store	upgrade to latest	3 weeks ago
LICENSE	upgrade to latest	3 weeks ago
README.md	update README	8 hours ago
docker-compose.yml	bump to latest	3 weeks ago

Step-by-Step Code Execution Instructions:

Quick Setup Option

- Make sure you have already installed docker (https://docs.docker.com/get-docker/) and docker-compose (https://docs.docker.com/compose/)
- Clone the Repository

```
$ git clone https://github.com/rohitc5/intel-oneAPI/
$ cd intel-oneAPI
```

- Start the LEAP RESTFul Service to consume both components (Ask Question/Doubt and Interactive Conversational AI Examiner) as a REST API. Also Start the webapp demo build using streamlit.

```
# copy the dataset
$ cp -r ./dataset webapp/

# build using docker compose
$ docker-compose build

# start the services
$ docker-compose up
```

- Go to http://localhost:8502

Model Checkpoint Release

https://huggingface.co/rohitsroch

H2S

intel

oneAPI

<HACK>ATHON

THANK YOU