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)

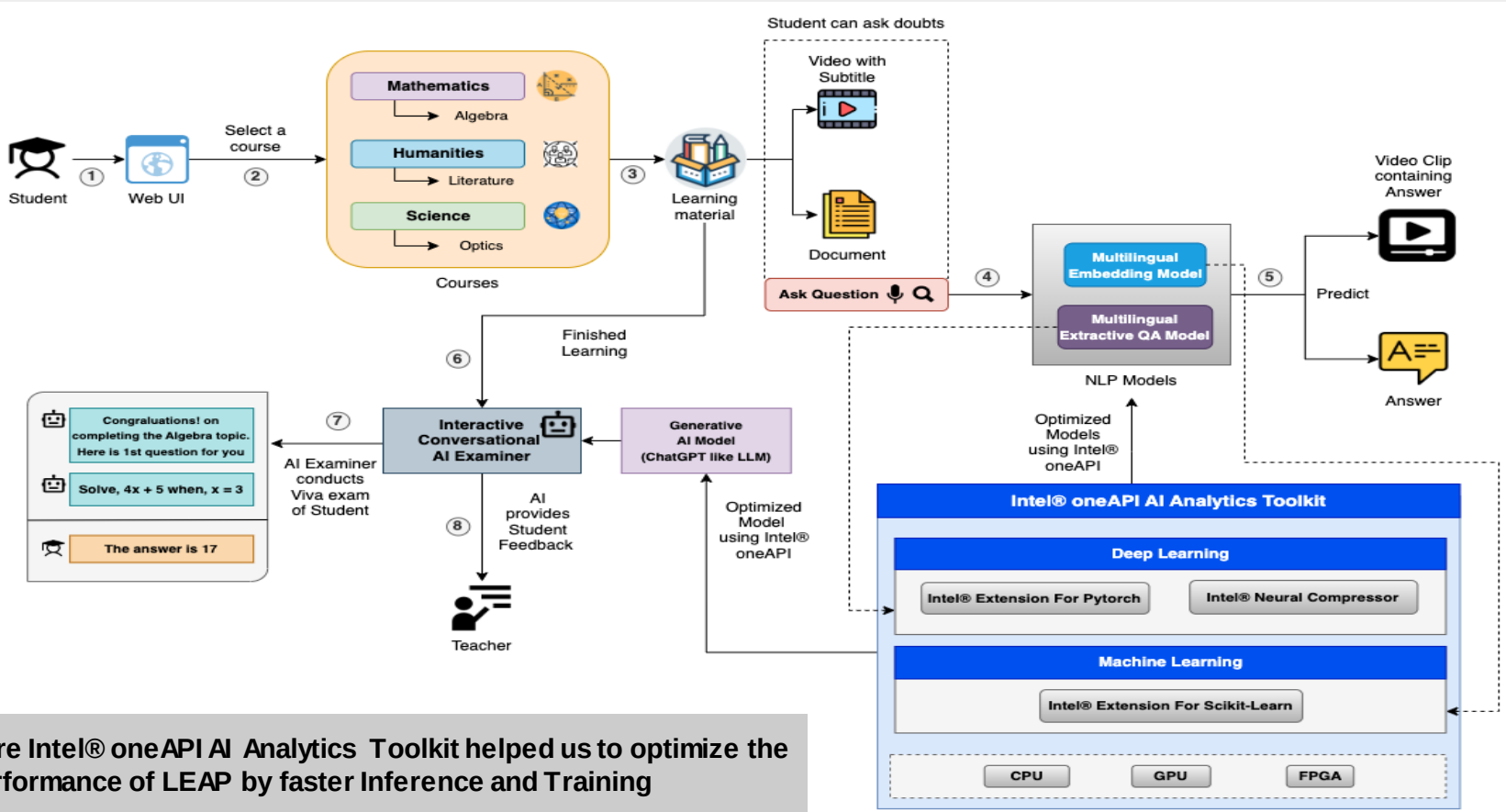
- AI based platform
- Powered by Intel OneAPI
- Quality Education
- All time Availability

Key Features of LEAP

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

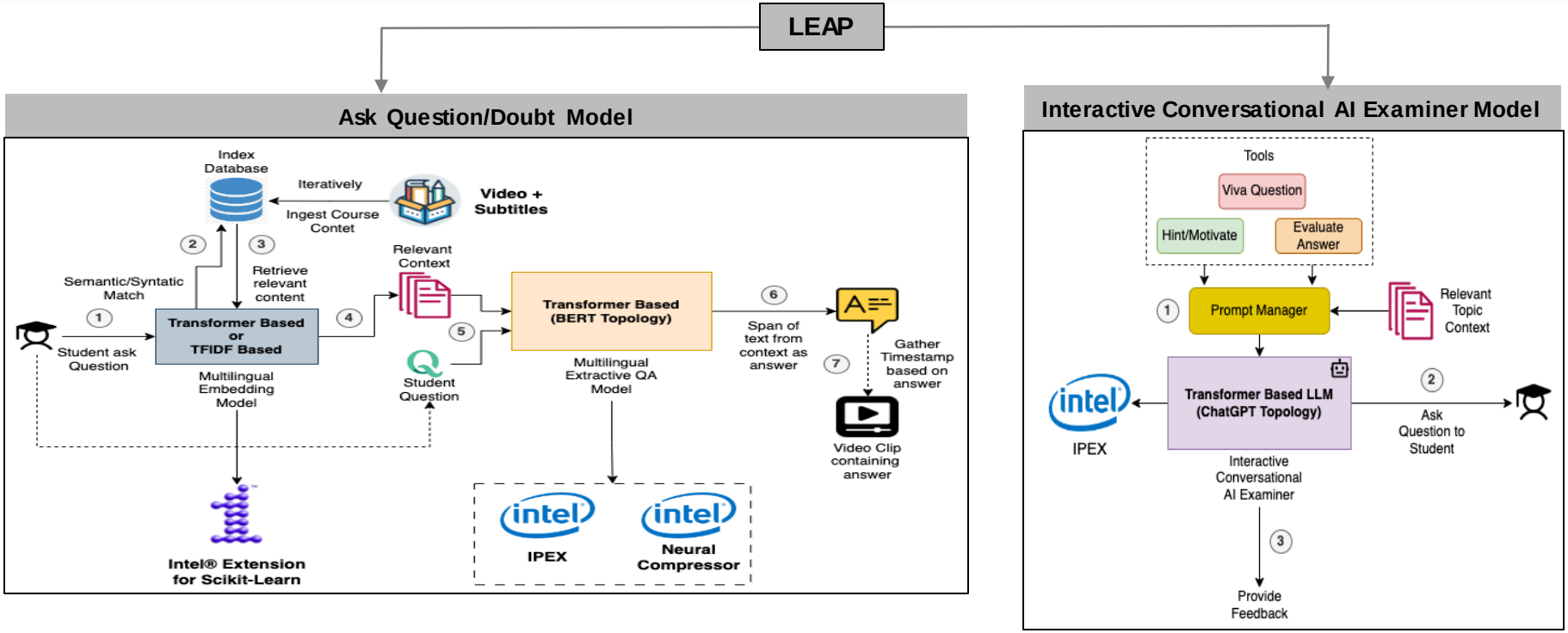
Reference: PRNewswire , Edtechreview; holoniq

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



Static-QAT-INT8 is 4.44x times faster in performance as compared to baseline For our Extractive QA Model

Extractive QA Model (BERT Topology) and TFIDFVectorizer's Latency, Throughput and F1 Score Comparison with IPEX and Intel® Neural Compressor

Ask Question/Doubt Model

Extractive Question Answering Model

	Pytorch (Base) - FP32	Pytorch (IPEX) - FP32	**Static-QAT-INT8**	Static-Smooth-QAT-INT8
Latency (milli sec)	64.513	39.329	**14.514**	15.24
Throughput (samples/sec)	15.501	25.427	**68.9**	65.616
F1 Score (SQuAD-v1)	**76.11**	**76.11**	75.72	75.72

Interactive Conversational AI Examiner Model

TFIDF Embedding Model

	Scikit-Learn (Base)	**Intel Extension For Scikit-Learn**
Latency (milli sec)	0.761	**0.752**
Throughput (samples/sec)	1313.63	**1330.49**

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

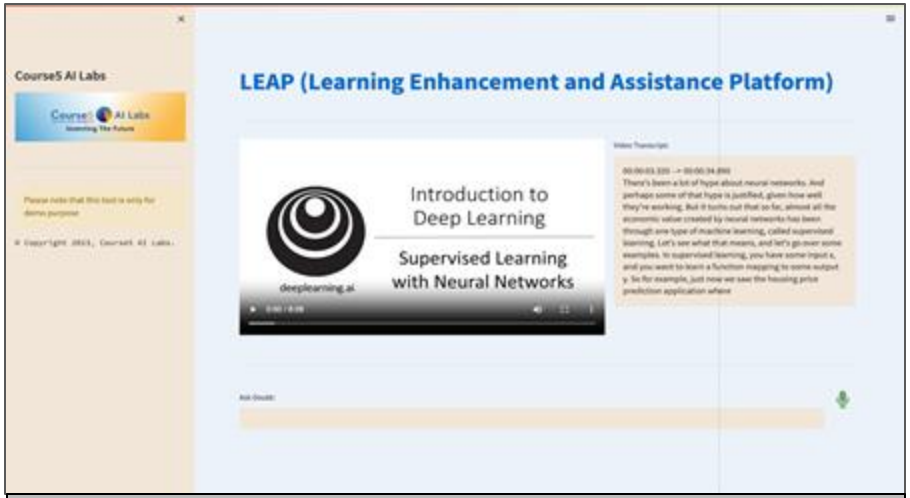
Demo Link and Screenshots

Link: https://www.youtube.com/watch?v=M51BFcoJa3k

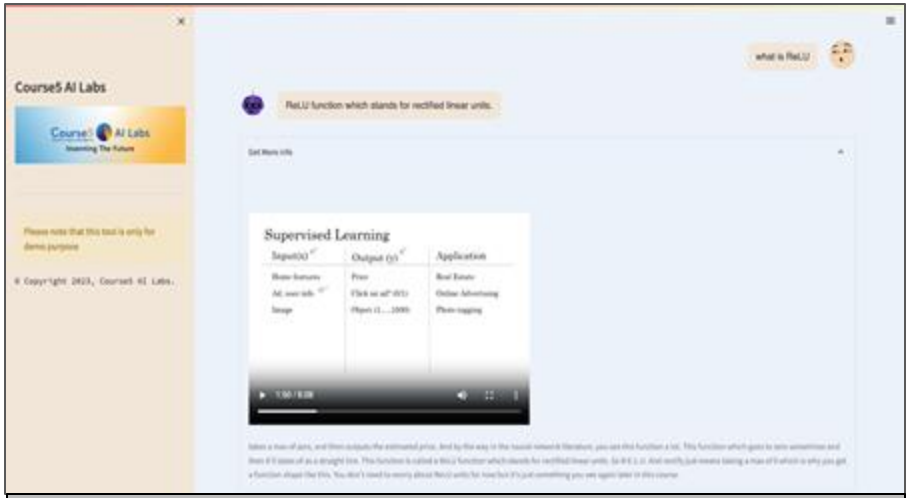
Screen-1 (Dashboard Page)

Screen-2 (Week wise Content Page)

Demo Screenshots



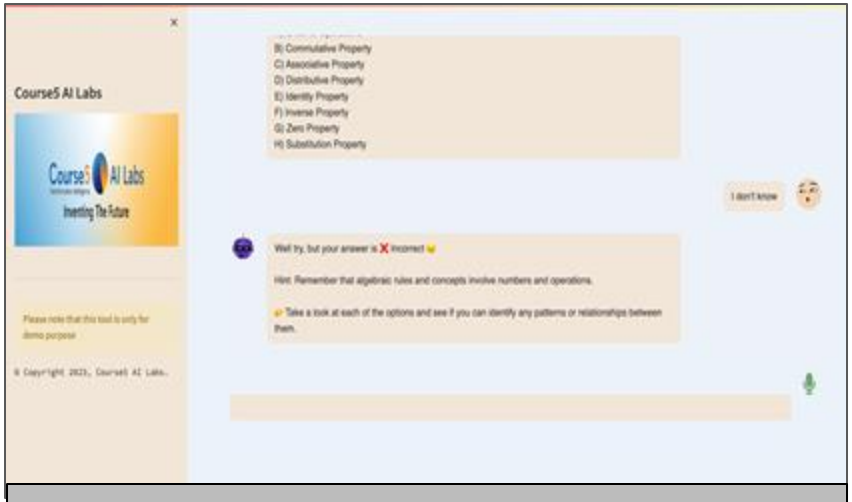
Screen-3 (Ask Question/Doubt Page)



Screen-4 (Ask Question/Doubt Page)

Demo Screenshots

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

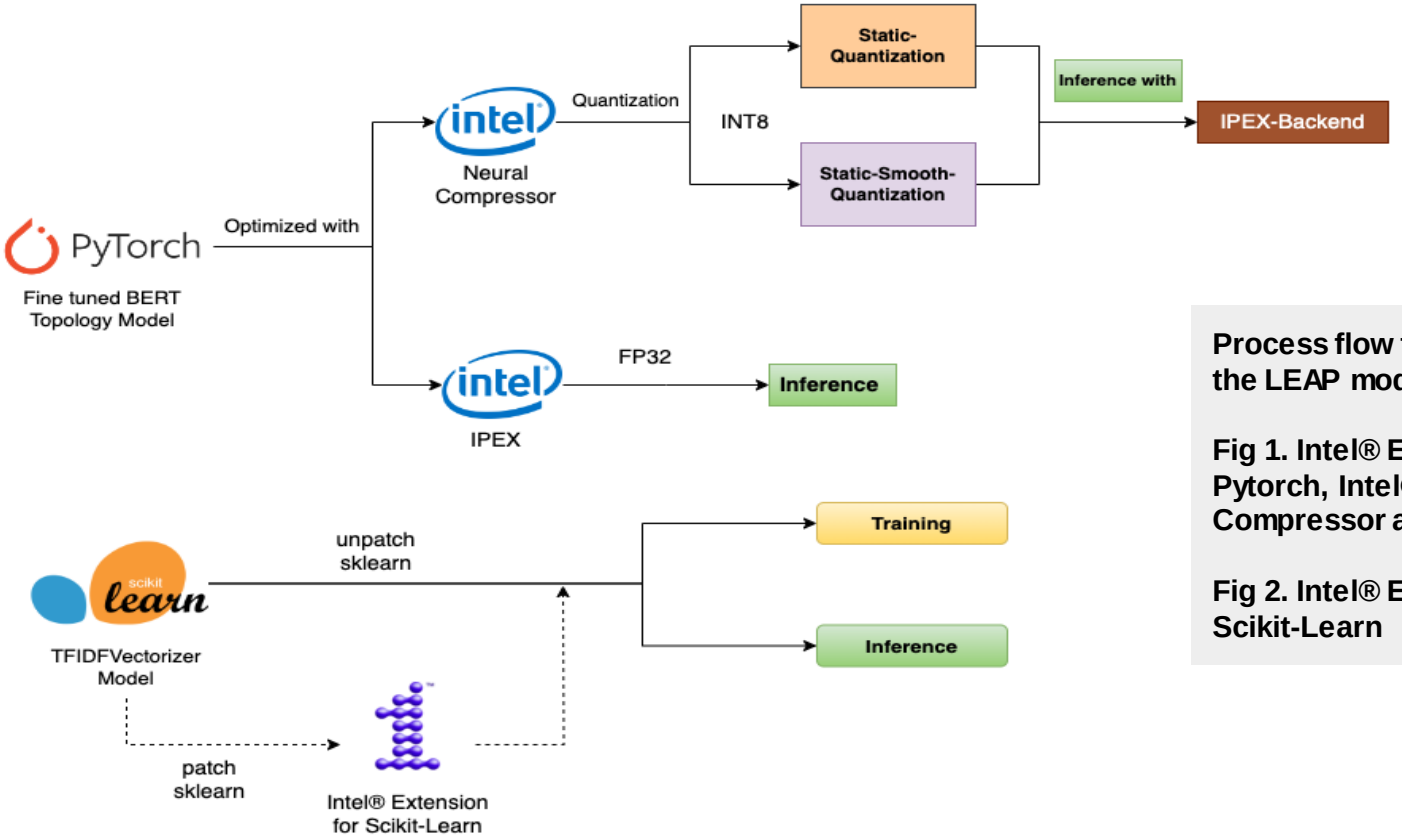


Screen-6 (Interactive Conversational AI Examiner provides hints and motivates a student in case of a wrong answer)

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

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

Result Summary (unique aspects of oneAPI/SYCL used)

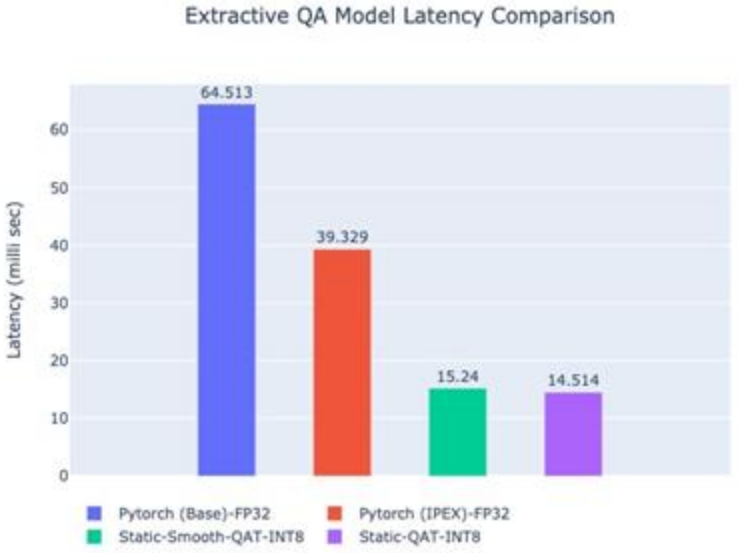


Process flow to optimize the LEAP models by using

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

Fig 2. Intel® Extension for Scikit-Learn

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



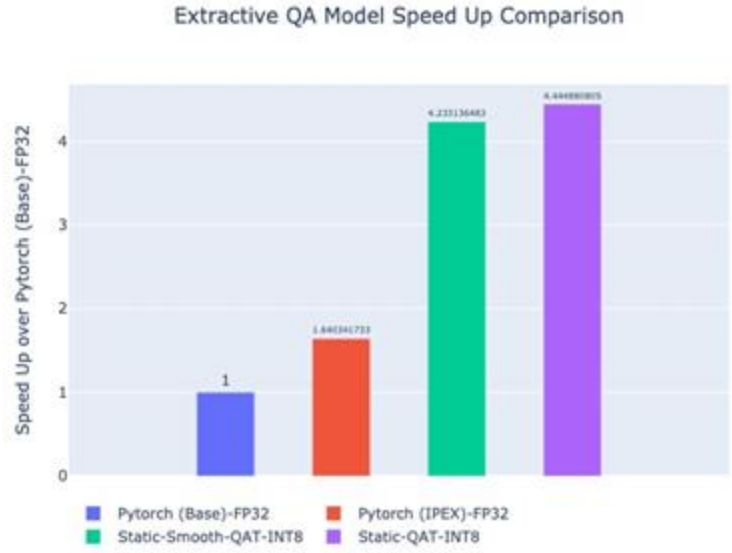
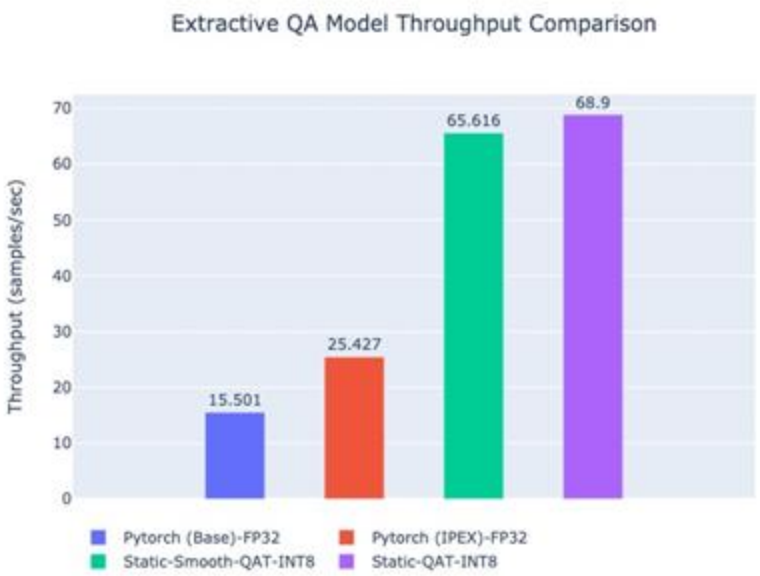


*Fig: Latency/Speed-Up Benchmark result for **our Extractive Question Answering ALBERT Model (Multilingual)** on Intel® Dev Cloud machine (**Intel® Xeon® Platinum 8480+ (4th Gen: Sapphire Rapids) - 224v CPUs 503GB RAM**) with optimization using IPEX-FP32 and Static INT8-Quantization using Intel® Neural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

For Ask Question/Doubt Extractive QA Model

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



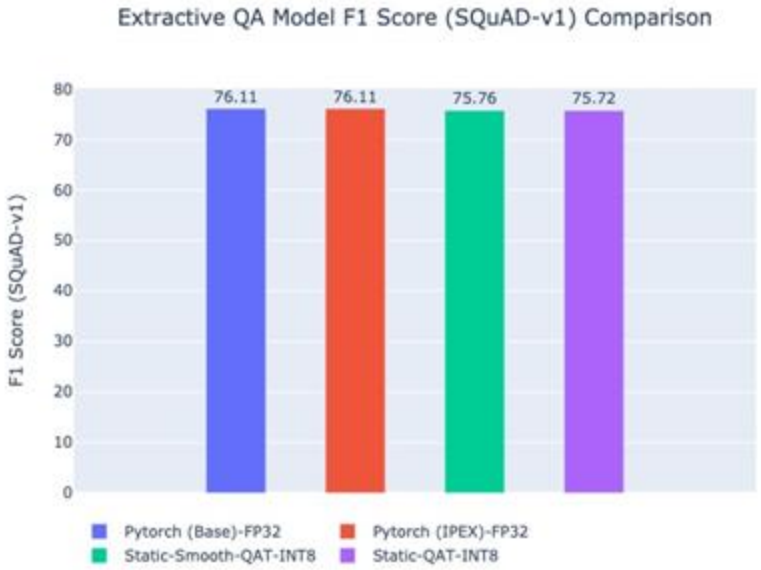


*Fig: Throughput/F1 Score Benchmark result for **our Extractive Question Answering ALBERT Model (Multilingual)** on Intel® Dev Cloud machine (**Intel® Xeon® Platinum 8480+ (4th Gen: Sapphire Rapids) - 224v CPUs 503GB RAM**) with optimization using IPEX-FP32 and Static INT8-Quantization using Intel® Neural 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

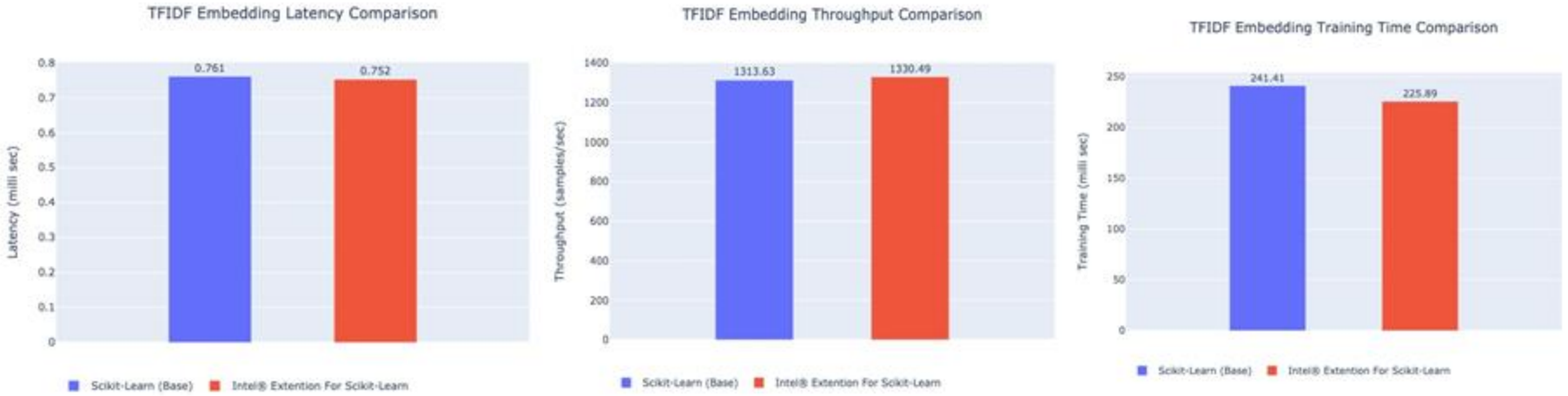


*Fig: Benchmark results for **TFIDFVectorizer** Embedding model during training and inference on Intel® Dev Cloud machine (**Intel® Xeon® Platinum 8480+ (4th Gen: Sapphire Rapids) - 224v CPUs 503GB RAM**). Please Note that we don't see much of a difference may be because we used a tiny dataset.*

For Ask Question/Doubt Embedding Model

https://huggingface.co/rohitsroch

H2S

intel

oneAPI

<HACK>ATHON

THANK YOU

Problem Statement

MOOCs (Massive Open Online Courses) have surged in popularity in recent years, particularly during the COVID-19 pandemic. These online courses are typically free or low-cost, making education more accessible worldwide.

Key Challenges

Online learning has become imperative to students. However, learning experience is not optimal, due to key challenges include:

1. To **sift through pile of lengthy videos** or documents to find relevant information
2. **Resolving doubts** can be a **time-consuming process**
3. **24x7 Teacher availability** for guidance

Objective

To mitigate challenges, we propose our **LEAP** (Learning Enhancement and Assistance Platform), which is an AI-powered platform designed to enhance student learning outcomes and provide equitable access to quality education.

Key Features

- Ask Question/Doubt
- Conversational AI Examiner
- Feedback from AI Examiner
- Highly Reliable