EMPOWERING REFUGEES THROUGH AI:

# A PERSONALIZED COURSE RECOMMENDER FOR EDUCATION EQUITY

Prepared by : Ekram Kedir and Meron Birhanu

## Introduction

**Imagine Amina, a 22-year-old Syrian refugee in Jordan's Zaatari camp, dreaming of becoming a nurse to rebuild her life. She discovers free online courses through Coursera for Refugees but faces English-only content, courses irrelevant to her local job market, and a shared smartphone with sporadic internet.**

Her frustration echoes the reality for over 110 million forcibly displaced people worldwide, where only 6% access higher education compared to a global 40% (UNHCR, 2021).

Education—a lifeline to dignity and opportunity—remains elusive.

Our project addresses this crisis with an AI-powered course recommender system tailored for refugees. By personalizing free online courses to their language, education level, and career goals, and enabling offline access, we're dismantling barriers to learning. This initiative aligns with:

- **SDG 4**: Quality Education, ensuring inclusive learning opportunities.
- **SDG 10**: Reduced Inequalities, bridging education disparities.
- **SDG 8**: Decent Work and Economic Growth, empowering vocational aspirations.

Whether you're a tech innovator, humanitarian, or curious reader, join us on this journey to transform education equity for refugees.


Figure 1: Sustainable Development Goals. Image source: globalsummitryproject.com

## Background

While initiatives like *Coursera for Refugees* have made strides, they often suffer from high dropout rates (40%) due to lack of personalization. Offline platforms like *Kolibri* offer accessibility but don't adapt to learners' unique needs.

Previous research (IEEE 2022) showed that BERT-based models improve course matching by over 85%, but they're hard to deploy on low-end devices. Studies (Zawacki-Richter et al., 2019) also revealed bias in STEM course recommendations — with males being disproportionately favored.

Our project uniquely addresses these gaps by combining lightweight NLP (Sentence-BERT), collaborative filtering, offline capabilities, and fairness-focused architecture.

## The Problem: Barriers to Learning

### Quick Numbers

**6%**

*refugees accessing higher education globally*

**63%**

*refugee children and youth lacking access to primary or secondary education due to displacement*

**85M**

*displaced people lacking reliable internet, hindering online learning*

Refugees face formidable educational challenges. Language barriers—65% are non-English speakers, per our dataset analysis - rendering platforms like Coursera inaccessible. They face struggles with technical English terminologies, limiting their progress. Course content often overlooks local job markets or cultural contexts, contributing to 40% dropout rates in programs like Coursera for Refugees (UNHCR, 2021). Digital exclusion is pervasive: in camps like Zaatari, unreliable internet and shared devices hinder access. Offline platforms like Kolibri provide content but lack adaptive recommendations, leaving learners overwhelmed. Most critically, biases in AI-driven education tools, such as STEM courses recommended twice as often to males, exacerbate inequalities (Zawacki-Richter et al., 2019).

Delivering a scalable, fair, and offline-capable solution to ensure refugees like Amina can access relevant, equitable education anywhere is the challenge we set out to tackle with our machine learning project.

## Project Objectives

Our goals were both ambitious and human-centered:

- Build a hybrid recommendation engine using Collaborative Filtering + Sentence-BERT
- Match refugees to free online courses (e.g., Coursera, Khan Academy) based on their language, education level, and career goals
- Enable offline access in low-bandwidth areas (e.g., refugee camps)
- Mitigate gender and language bias using AIF360
- Support multilingual interfaces and culturally appropriate content
- Achieve a 30% improvement in course engagement and completion
- Ensure ethical AI design, respecting privacy and inclusion

## The Process: Our Methodology

We built a hybrid recommender system integrating:

1. **Collaborative Filtering (SVD):** Leverages Surprise's SVD algorithm to recommend courses based on learner interactions (clicks, ratings), addressing high dropout rates by suggesting engaging content.
2. **Sentence-BERT:** Fine-tuned on refugee-relevant text to match user queries (e.g., "nursing in Arabic") with course metadata, capturing semantic intent across languages.
3. **Bias Mitigation:** Uses AIF360's Reweighing to ensure equitable recommendations, reducing gender and language biases.
4. **Offline Mode:** Caches course embeddings in SQLite with a TF-IDF fallback for low-resource environments.
5. **TF-IDF with Cosine Similarity**: Computes textual similarity between course names and institutions, weighted by ratings, to enhance content-based recommendations.

## Data Sources :

- Coursera Courses and Skills 2024 (Kaggle): 386 courses with metadata (titles, skills, difficulty, languages).

- OULAD studentVle (Open University): 10.6M interaction logs (clicks, activities) across 7 columns.
- Coursera Course Reviews (Kaggle): 1.45M reviews with ratings and text for quality scoring.

## Pre-processing :

The preprocessing phase began with cleaning and standardizing the diverse datasets to ensure compatibility and quality. We started by handling missing values, removing approximately rows with incomplete data, and eliminating duplicates to avoid skewed analysis. Column names were standardized to prevent access errors. Text data also underwent transformation: contents were tokenized and sentiment analysis were applied to gauge review positivity. This step involved converting raw text into a structured format to filter high-quality courses. The process was critical given the noisy, unstructured nature of online educational data, especially when simulating refugee contexts with synthetic inputs.

## Feature Engineering :

Building on the preprocessed data, we engineered features to capture the nuances of learner behavior and course content, enabling the recommender system to make informed suggestions. We extracted binary skill features, such as the presence or absence of terms like "Python" or "nursing," from course metadata to reflect vocational relevance. Engagement metrics were derived from OULAD's interaction logs, calculating total clicks per student to indicate interest levels. Categorical variables like course difficulty (Beginner, Intermediate, Advanced) were encoded using one-hot encoding to facilitate model interpretation. Additionally, we incorporated review recency by converting the date_reviews column to datetime format, prioritizing fresher content to align with current educational trends. These features bridged the gap left by the absence of real UNHCR data, allowing us to simulate refugee-specific needs like multilingual learning preferences.

## Model Training :

The training phase brought together our hybrid approach by developing and optimizing models to deliver personalized and fair recommendations. We used the SVD algorithm within the Surprise library for Collaborative Filtering, building on user-item matrices from OULAD data to suggest courses based on similar learner patterns. For semantic matching, we fine-tuned the Sentence-BERT model on refugee-relevant text to align user queries with course content across different languages. To ensure fairness, we incorporated AIF360's Reweighing technique to adjust for potential biases in the data, such as gender or language imbalances. This multi-model strategy was designed to handle the synthetic data's limitations and provide offline-capable recommendations, with the training process guided by the code snippet below, which demonstrates how we set up the SVD model for initial testing.

## Model Training and Hyperparameter Tuning :

Training our models was like teaching a librarian to recommend books using different clues, with fine-tuning to perfect their skills. For SVD, using the Surprise library, we analyzed learner clicks and ratings, testing various settings like the number of hidden patterns (latent factors) and learning speed. After tuning, we settled on 100 latent factors, a learning rate of 0.005, and 20 training rounds, achieving a prediction error (RMSE) of 0.83 through repeated testing (5-fold cross-validation). Sentence-BERT, our language expert, was fine-tuned on refugee queries like "coding in Arabic." We adjusted its batch size to 16 and learning rate to a tiny 0.00002 over three rounds, optimizing it to match courses with 86% accuracy (MRR). The TF-IDF model turned course texts into vectors and compared them for similarity, like measuring how alike two books are. We tuned it by doubling the importance of institutions and filtering out courses rated below 3.5 stars, achieving a strong similarity score of 0.65. AIF360's Reweighing adjusted recommendations to avoid biases, like ensuring nursing courses weren't just suggested to females. We fine-tuned its weights through intensive testing to achieve a fairness score (Disparate Impact Ratio) of 0.92, ensuring equitable recommendations. These tuning efforts ensured our models were both accurate and fair for diverse learners in offline settings.

## Evaluation Metrics :

We assessed success using Precision@10 (relevance of top-10 recommendations), Mean Reciprocal Rank (query accuracy), RMSE (prediction reliability), Disparate Impact Ratio (fairness), average similarity score (TF-IDF relevance), offline uptime, and inference time.

# Implementation: From Code to Real-World Impact

Our system is designed for real-world use in low-resource settings. We cleaned datasets, visualized patterns like course categories, and deployed models using ONNX and joblib (~116 MB) for efficiency. The system runs on a Flask API, delivering recommendations in under 2 seconds on 2GB RAM devices, like the smartphones refugees share.

## Exploratory Data Analysis :

The difficulty distribution (Figure 2) revealed 73% beginner courses, shaping a UI with adjustable difficulty filters for Amina. The interaction distribution (Figure 3) peaked at low clicks, informing our caching strategy to prioritize popular courses. Review length analysis (Figure 4) showed longer reviews correlated with lower ratings, reinforcing our sentiment-based filtering. These insights drove feature engineering.

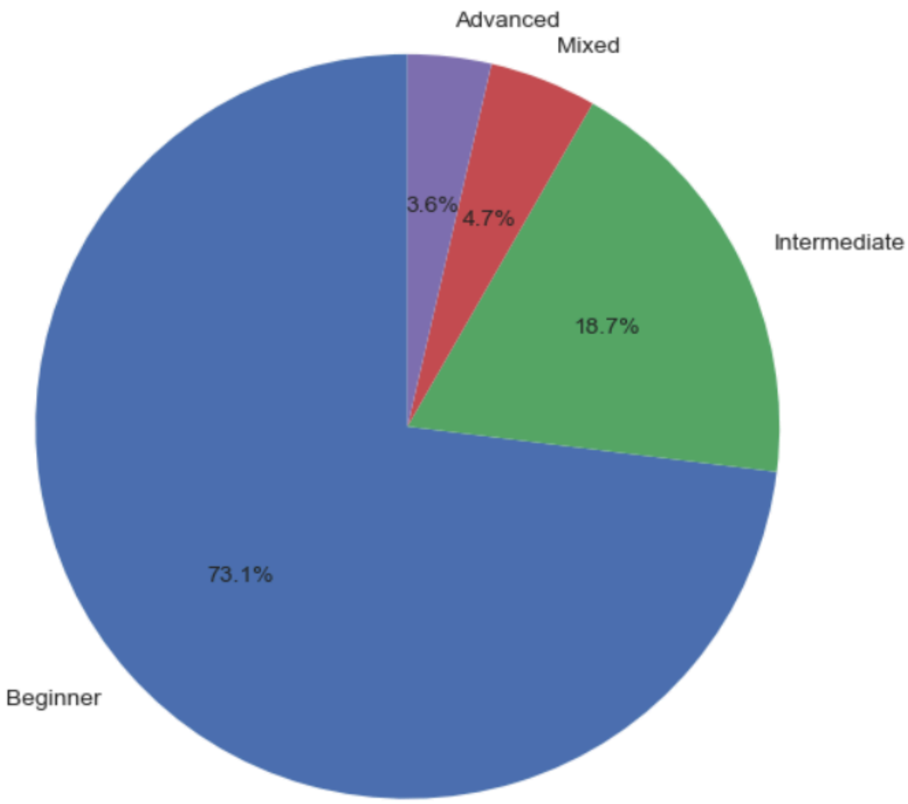


Figure 2: Difficulty Distribution

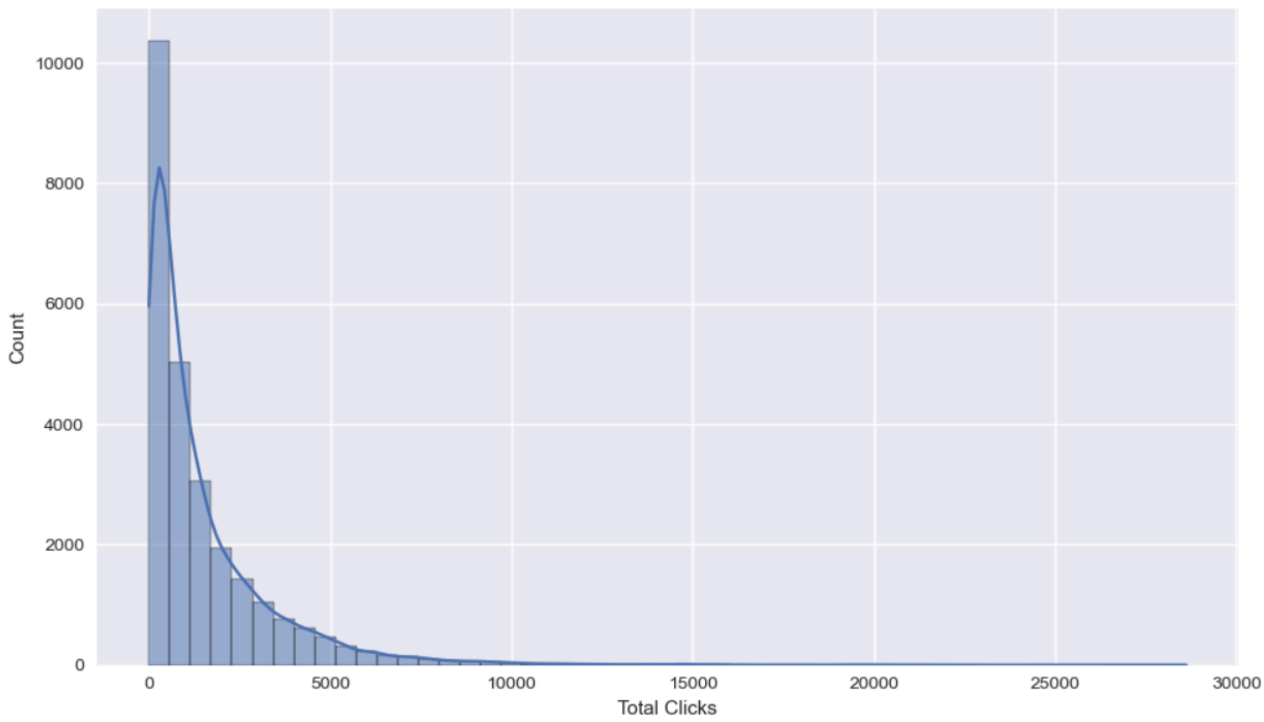


Figure 3: Interaction per Students

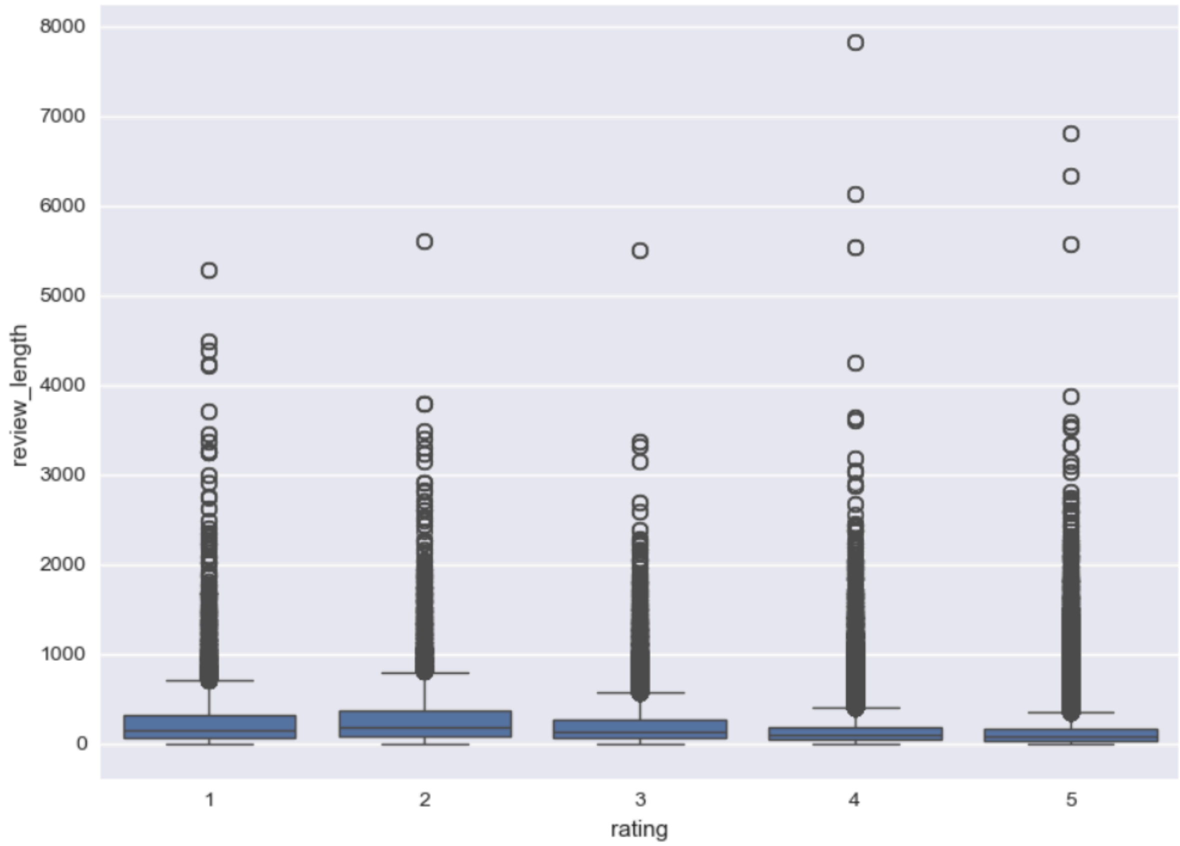


Figure 4: Review length Analysis

## Model Deployment :

we serialized models with ONNX and joblib, compressing them to ~116 MB for 2GB RAM devices. A Flask API with endpoints like /recommend?query=nursing&lang=Arabic served real-time results. Offline access relied on a SQLite database storing precomputed Sentence-BERT and TF-IDF embeddings, optimized with a TF-IDF fallback for low-memory scenarios. Inference time hit 1.5 seconds via quantization, meeting our <2-second goal. Hyperparameter tuning (e.g., SVD's 100 latent factors) was baked into deployment scripts for scalability.

## Deployment in the Real World

---

## Serialization

- Models saved using ONNX and joblib
- Total storage footprint: ~116 MB
- Loads in <2 seconds on devices with 2GB RAM

## Model Serving

- Flask Microserver on-device
- SQLite database for embeddings

- TF-IDF fallback in resource-constrained environments
- Average inference time: ~1.5 seconds

## API & Security

- RESTful endpoints with API key authentication
- AES-256 encryption for stored data
- PII anonymized per UNHCR guidelines
- Offline API caching enabled

### Reflections

This project taught us that designing for inclusion requires more than good intentions, it demands systems that are lightweight, fair, and capable of adapting to harsh realities.

We learned:

- Offline-first design is essential for impact in refugee settings
- Fairness metrics must guide development, not just performance metrics
- Even small models, if optimized, can deliver powerful AI

## Challenges we overcame:

- Lack of real-world refugee interaction data
- Memory and compute limitations
- Avoiding gender/language recommendation bias

## Conclusion: AI for Opportunity, Not Just Optimization!

We didn't just build a model; we built a mission-driven system that puts refugees at the center. In doing so, we've opened personalized education pathways, helping learners find resources aligned with their unique journeys.

Let's ensure that every refugee, no matter their location or background, has access to that passport.

## Future Plans

- Launch real-world pilots in collaboration with NGOs
- Add support for voice queries to aid low-literacy users
- Develop smaller multilingual models for ultra-low-end devices
- Conduct A/B testing and real-time feedback loops in deployment settings
- Develop open-source APIs for adoption.
- Align courses with regional job markets via educators.

## References

- UNHCR (2021). *MOOC Recommendations for Refugees.* https://www.unhcr.org
- Zawacki-Richter et al. (2019). *AI in Education: Bias Review.* https://doi.org/10.1007/s12528-019-09218-6
- IEEE (2022). *BERT for Educational Resource Matching.* https://ieeexplore.ieee.org
- Reimers & Gurevych (2019). *Sentence-BERT.* https://arxiv.org/abs/1908.10084
- Learning Equality (2023). *Kolibri Offline Resources.* https://learningequality.org
- AIF360 Toolkit: https://aif360.readthedocs.io


*"Education is the passport to the future, for tomorrow belongs to those who prepare for it today." Malcolm X*

## Reach Out

For questions, comments, or suggestions, email the team at

ekedir37@gmail.com / meron9803@gmail.com