INVENTION DISCLOSURE FORM

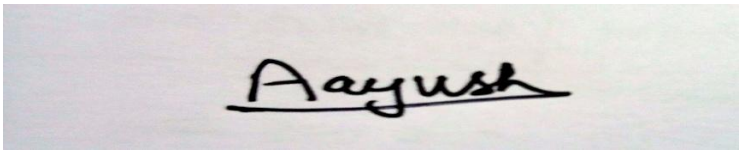
Annexure–3B: Complete Filing

1. TITLE

AI-Based Job and Course Recommendation System Using TF-IDF, Machine Learning Evaluation, and Cosine Similarity Ranking

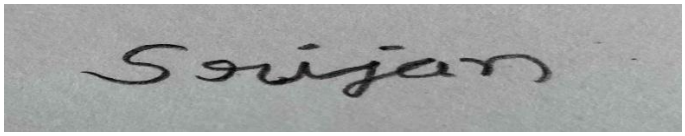
2. INTERNAL INVENTOR(S) / STUDENT(S)

A. Inventor 1

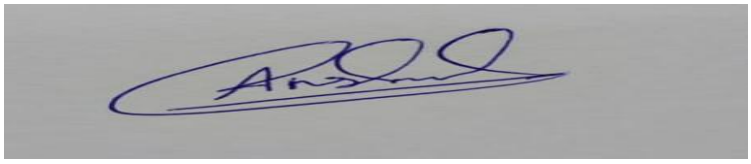
Full Name	Aayush Sharma
Mobile Number	9887075729
Email (Personal)	aayushsharma1715@gmail.com
Registration Number	12317299
Address	Lovely Professional University, Punjab–144411, India
Signature	

B. Inventor 2

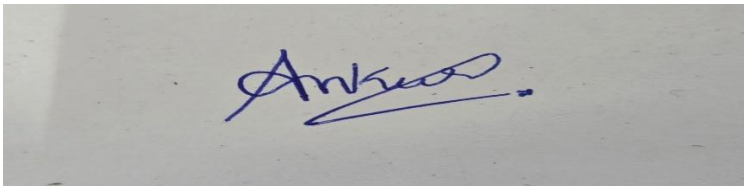
Full Name	Srijan Kumar Pandey
Mobile Number	9343473399
Email (Personal)	srijanp56@gmail.com
Registration Number	12316351
Address	Lovely Professional University, Punjab–144411, India

Signature	

C. Inventor 3

Full Name	Anshul Bhuria
Mobile Number	8628840708
Email (Personal)	Anshulbhuria2006@gmail.com
Registration Number	12324584
Address	Lovely Professional University, Punjab–144411, India
Signature	

D. Inventor 4

Full Name	Ankur
Mobile Number	9131898526
Email (Personal)	ankurthemewtwo@gmail.com
Registration Number	12318814
Address	Lovely Professional University, Punjab–144411, India
Signature	

3. DESCRIPTION OF THE INVENTION

A. PROBLEM ADDRESSED BY THE INVENTION

The modern learning environment offers millions of online courses and thousands of job postings, yet no unified system helps learners understand:

- Which job roles match their current skills,
- Which skills they are missing for desired job roles,
- Which online courses best fill these skill gaps,
- Which machine learning model performs best for classification,
- How to generate accurate, personalized job and course recommendations.

Existing platforms operate independently—job portals list vacancies without skill roadmaps, and course platforms offer training without job relevance.

Manual matching of skills to jobs is impractical, subjective, and incomplete.

The proposed AI system resolves these issues by:

- Processing job and course datasets using TF-IDF vectorization,
- Generating synthetic labels to enable supervised ML training,
- Training and evaluating multiple ML models (KNN, SVM, Logistic Regression),
- Automatically selecting the best model,
- Recommending jobs and courses using cosine similarity ranking,
- Creating a unified, automated skill–career–learning mapping system.

B. OBJECTIVES OF THE INVENTION

1. To automatically classify job and course descriptions using TF-IDF embedding.
2. To generate synthetic labels for supervised ML training from unstructured text.
3. To train and evaluate KNN, SVM, and Logistic Regression models.
4. To create a ranking-based recommendation engine using cosine similarity.
5. To recommend job roles that best match user skills.
6. To identify missing skills and suggest the most relevant online courses.
7. To develop a scalable Python backend deployable with React frontend.

C. STATE OF THE ART / RESEARCH GAP / NOVELTY

Sr	Patent ID / Prior Work	Abstract Summary	Research Gap	Novelty in This Invention
1	US20160232043A	1 Basic text classification into sentiment categories.	No job–course integration, no skill mapping, no model comparison.	Unified job–course recommendation system + ML-model evaluation + cosine similarity engine.
2	US20190291721A	1 Deep learning for emotion detection.	Focuses only on classification, not job-role prediction or skill gaps.	Synthetic labeling + job-role clustering + course gap mapping.

3	US20200345612A	1 Toxicity detection using neural networks.	Not applicable for career or course recommendation.	Career-path mapping, skill-gap analysis, dual-dataset recommendation approach.

Novelty Summary:

- First hybrid system combining ML classification with cosine similarity for job and course recommendations.
- Synthetic category labeling from raw text to enable supervised training without manual labels.
- Automatic best-model selection based on accuracy metrics.
- Unified mapping of skills → job roles → suitable online courses.

D. DETAILED DESCRIPTION

1. Overview

The system loads job and course datasets, preprocesses text, generates TF-IDF vectors, creates synthetic labels, trains ML models, evaluates performance, selects the best model, and finally uses cosine similarity to recommend job roles and courses.

2. System Requirements

Hardware Requirements

- Minimum: Dual-Core CPU, 6GB RAM, 10–20GB Storage
- Recommended: Intel i5/Ryzen 5, 8GB+ RAM, 50GB+ Storage

Software Requirements

- Python 3.8+
- Libraries: pandas, numpy, scikit-learn, nltk, joblib
- Frontend: React.js, Bootstrap, Chart.js
- Backend: Flask / FastAPI

3. **Data Sources**
   - Naukri Job Dataset (CSV)
   - Udemy Course Dataset (CSV)
4. **Data Processing and Feature Engineering**
   - Cleaning text, filling missing fields
   - Combining job title + skills + description
   - Combining course title + subject
   - Lowercasing, token normalization
   - TF-IDF vectorization for numerical feature extraction
5. **Machine Learning Modules**

   a) Synthetic Labeling

      - Data Science (Category 0)
      - Software Development (Category 1)
      - Design (Category 2)
      - Business (Category 3)

- Others (Category 4)

b) Model Training

• KNN Classifier

•SVM Classifier

• Logistic Regression

c) Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1-score

6. Cosine Similarity Recommendation Engine

- User-skill vector vs job TF-IDF matrix → Top job matches
- User-skill vector vs course TF-IDF matrix → Top course recommendations

7. Working Prototype

A functional prototype is deployed using Python backend with React frontend, supporting:

- Job prediction
- Course recommendation
- Model performance reporting

- Skill gap analysis

4. RESULTS AND ADVANTAGES

Results

- Accurate job-role clustering using synthetic labeling.
- Measured model performance using four metrics.
- Effective ranking of job and course results using cosine similarity.

Advantages

1. Automates job and course matching.
2. Provides transparent ML model evaluation.
3. Scales easily for large datasets.
4. Works without manually labeled datasets.
5. Provides personalized, data-driven recommendations.

5. USE AND DISCLOSURE

A. Have you described or shown your invention/ design to anyone or in any conference?	YES ( )	NO (X)
B. Have you made any attempts to commercialize your invention (for example, have you approached any companies about purchasing or manufacturing your invention)?	YES ( )	NO (X)

C. Has your invention been described in any printed publication, or any other form of media, such as the Internet?	YES ( )	NO (X)
D. Do you have any collaboration with any other institute or organization on the same? Provide name and other details.	YES ( )	NO (X)
E. Name of Regulatory body or any other approvals if required.	YES ( )	NO (X)

6. POTENTIAL COMMERCIALIZATION

- University placement departments
- Job portals and career guidance companies
- EdTech platforms for personalized course recommendation
- Corporate upskilling and hiring systems

7. COMPANIES FOR COMMERCIALIZATION

- Naukri.com – https://www.naukri.com
- Udemy Business – https://business.udemy.com
- Coursera – https://www.coursera.org
- LinkedIn Learning – https://www.linkedin.com/learning
- Scaler Academy – https://www.scaler.com
- UpGrad – https://www.upgrad.com

8. ROYALTY AND PATENT LIABILITY

No patented algorithms are used. TF-IDF, cosine similarity, and classical ML algorithms are public-domain mathematical techniques.

9. PATENT FILING OPTION

Recommended filing option: **Provisional Patent Application**.

10. FILLING OPTIONS

Provisional Patent Filing:

I am considering filing a provisional patent to establish an early priority date for my invention. Since I have a well-defined algorithm and a basic working prototype, this filing will give me 12 months to further refine and develop the technology before converting it into a complete patent. This approach is ideal for me because my AI model, skin diagnostics method, and allergy detection process are unique but still evolving.

11. KEYWORDS

Machine Learning, TF-IDF, Cosine Similarity, Job Recommendation, Course Recommendation, Skill Gap Analysis, SVM, KNN, Logistic Regression.

(Letter Head of the external organization)

NO OBJECTION CERTIFICATE

This is to certify that University/Organization Name or its associates shall have no objection if Lovely Professional University files an IPR (Patent/Copyright/Design/any other…….) entitled "……." including the name(s) of, ...as inventors who is(are) student(s)/employee(s) studying/ working in our University/ organization.
Further Name of the University/Organization shall not provide any financial assistance in respect of said IPR nor shall raise any objection later with respect to filing or commercialization of the said IPR or otherwise claim any right to the patent/invention at any stage.

(Authorised Signatory)