

Project name: #1 Personalized E-Learning Recommendation System

Problem statement: In online education platforms (Coursera, Udemy, etc.), students get overwhelmed with too many courses. A personalized recommendation engine helps suggest the right courses based on student interests and learning history.

What was the problem:

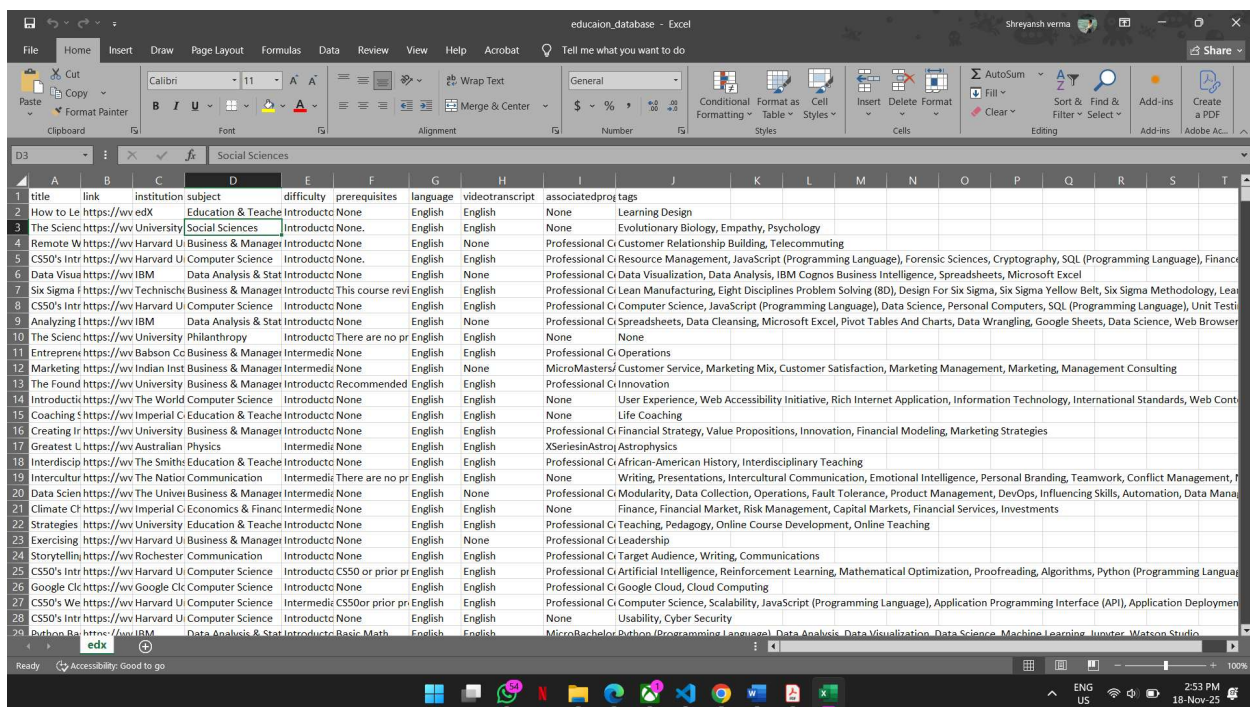
A personalized e-learning app is essential today because learners face overwhelming amounts of information and diverse skill demands. Traditional, one-size-fits-all learning often fails to meet individual needs, pace, or interests. A personalized platform adapts content to each learner's background, goals, and performance, making learning more efficient and engaging. It offers tailored recommendations, flexible learning paths, and instant feedback, helping users stay motivated and improve faster. Additionally, it saves time, promotes continuous upskilling, and supports learners of different abilities. In a fast-changing digital world, personalized e-learning ensures that learning is relevant, accessible, and aligned with real-world needs.

Moreover, personalized e-learning apps leverage data analytics and AI to understand user behavior, strengths, and weaknesses, allowing them to deliver highly targeted content. This creates a more immersive and effective learning experience compared to traditional classroom methods. Such platforms support self-paced learning, enabling students and professionals to balance education with busy schedules. They also enhance retention by presenting lessons in preferred formats—videos, quizzes, summaries, or hands-on tasks. Personalized learning helps reduce dropout rates by keeping learners engaged and confident. As industries evolve rapidly, these apps ensure individuals stay competitive by continuously adapting learning materials to match modern skill requirements.

Collection of Data:

Here I have used Kaggle to download a dataset which contains relevant information regarding this project and downloaded an excel file named 'education database' from <https://www.kaggle.com/search?q=Personalized+E-Learning+Recommendation+System>. The excel file of online courses dataset (Kaggle) contains features like subject, difficulty, ratings, tags which are picked up and processed by the code I wrote to display the information on the dashboard.

From time to time I got stuck in errors and loopholes to overcome them I had used Chat gpt <https://chatgpt.com/c/69133da7-f748-8321-a79f-d734660a5f9d>. this helped me in tackling challenges and clearing my doubts regarding a lot of things.



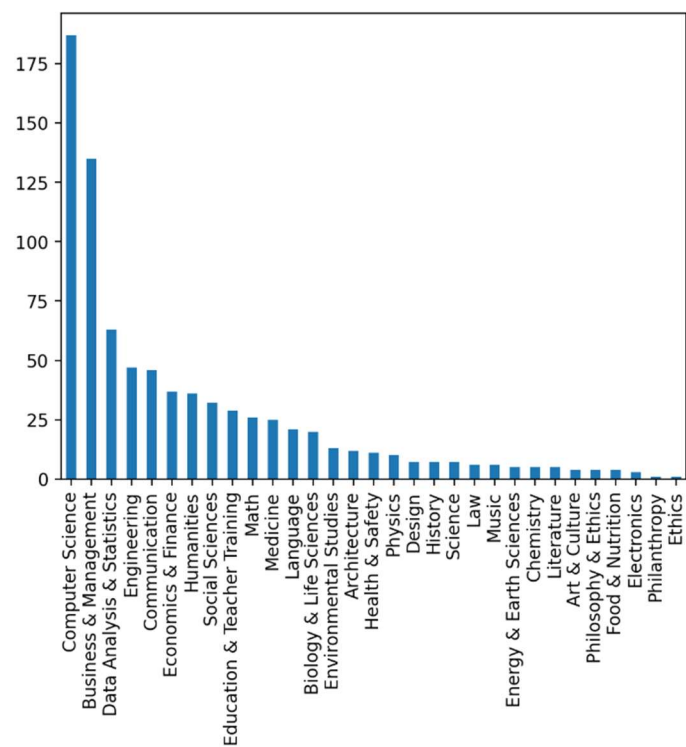
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	title	link	institution	subject	difficulty	prerequisites	language	videotranscript	associatedproj	tags										
2	How to Le	https://www.edx	Education & Teach	Introduct	None	English	English	None	None	Learning Design										
3	The Scien	https://www.university	Social Sciences	Introduct	None	English	English	None	None	Evolutionary Biology, Empathy, Psychology										
4	Remote W	https://www.harvard	Business & Manage	Introduct	None	English	English	None	Professional C	Customer Relationship Building, Telecommuting										
5	CS50's Intr	https://www.harvard	Computer Science	Introduct	None	English	English	None	Professional C	Resource Management, JavaScript (Programming Language), Forensic Sciences, Cryptography, SQL (Programming Language), Finance										
6	Data Visual	https://www.ibm	Data Analysis & Stat	Introduct	None	English	English	None	Professional C	Data Visualization, Data Analysis, IBM Cognos Business Intelligence, Spreadsheets, Microsoft Excel										
7	Six Sigma F	https://www.harvard	Business & Manage	Introduct	This course revi	English	English	None	Professional C	Lean Manufacturing, Eight Disciplines Problem Solving (8D), Design For Six Sigma, Six Sigma Yellow Belt, Six Sigma Methodology, Lea										
8	CS50's Intr	https://www.harvard	Computer Science	Introduct	None	English	English	None	Professional C	Computer Science, JavaScript (Programming Language), Data Science, Personal Computers, SQL (Programming Language), Unit Testi										
9	Analyzing I	https://www.ibm	Data Analysis & Stat	Introduct	None	English	English	None	Professional C	Spreadsheets, Data Cleansing, Microsoft Excel, Pivot Tables And Charts, Data Wrangling, Google Sheets, Data Science, Web Browser										
10	The Scien	https://www.university	Philanthropy	Introduct	There are no pr	English	English	None	None	None										
11	Entrepre	https://www.babson	Business & Manage	Intermedi	None	English	English	None	Professional C	Operations										
12	Marketing	https://www.indian	Business & Manage	Intermedi	None	English	English	None	MicroMasters	Customer Service, Marketing Mix, Customer Satisfaction, Marketing Management, Marketing, Management Consulting										
13	The Found	https://www.university	Business & Manage	Introduct	Recommended	English	English	None	Professional C	Innovation										
14	Introducti	https://www.the-world	Computer Science	Introduct	None	English	English	None	None	User Experience, Web Accessibility Initiative, Rich Internet Application, Information Technology, International Standards, Web Cont										
15	Coaching S	https://www.imperial	Education & Teach	Introduct	None	English	English	None	None	Life Coaching										
16	Creating Ir	https://www.university	Business & Manage	Introduct	None	English	English	None	Professional C	Financial Strategy, Value Propositions, Innovation, Financial Modeling, Marketing Strategies										
17	Greatest L	https://www.australian	Physics	Intermedi	None	English	English	None	XSeriesinAstro	Astrophysics										
18	Interdiscip	https://www.the-smiths	Education & Teach	Introduct	None	English	English	None	Professional C	African-American History, Interdisciplinary Teaching										
19	Intercultur	https://www.the-natio	Communication	Intermedi	There are no pr	English	English	None	None	Writing, Presentations, Intercultural Communication, Emotional Intelligence, Personal Branding, Teamwork, Conflict Management, f										
20	Data Scien	https://www.the-univel	Business & Manage	Intermedi	None	English	English	None	Professional C	Modularity, Data Collection, Operations, Fault Tolerance, Product Management, DevOps, Influencing Skills, Automation, Data Mana										
21	Climate Ch	https://www.imperial	Economics & Financ	Intermedi	None	English	English	None	None	Finance, Financial Market, Risk Management, Capital Markets, Financial Services, Investments										
22	Strategies	https://www.university	Education & Teach	Introduct	None	English	English	None	Professional C	Teaching, Pedagogy, Online Course Development, Online Teaching										
23	Exercising	https://www.harvard	Business & Manage	Introduct	None	English	English	None	Professional C	Leadership										
24	Storytelling	https://www.rochester	Communication	Introduct	None	English	English	None	Professional C	Target Audience, Writing, Communications										
25	CS50's Intr	https://www.harvard	Computer Science	Introduct	CS50 or prior pr	English	English	None	Professional C	Artificial Intelligence, Reinforcement Learning, Mathematical Optimization, Proofreading, Algorithms, Python (Programming Language)										
26	Google Clc	https://www.google	Computer Science	Introduct	None	English	English	None	Professional C	Google Cloud, Cloud Computing										
27	CS50's We	https://www.harvard	Computer Science	Intermedi	CS50or prior pr	English	English	None	Professional C	Computer Science, Scalability, JavaScript (Programming Language), Application Programming Interface (API), Application Deploymen										
28	CS50's Intr	https://www.harvard	Computer Science	Introduct	None	English	English	None	None	Usability, Cyber Security										
29	Python Ba	https://www.ibm	Data Analysis & Stat	Introduct	Basic Math	English	English	None	MicroBachelor	Python (Programming Language), Data Analysis, Data Visualization, Data Science, Machine Learning, Jupyter, Watson Studio										

Above is the snapshot of the Dataset used.

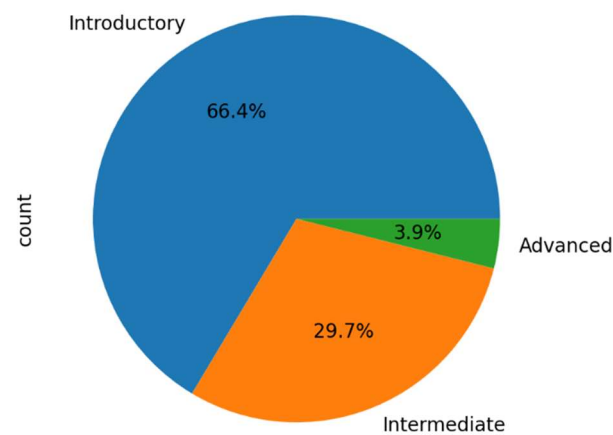
Data visualization:

The code visualizes the dataset using bar charts and pie charts to help users quickly understand key patterns within the available courses. For example, the bar chart displays the distribution of subjects, allowing learners to see which fields have the most content. The pie chart illustrates the difficulty levels across courses, giving users an overview of how beginner-friendly or advanced the platform is. These visualizations transform raw data into easy-to-interpret insights, making it simpler for users to explore trends without manually scanning large tables. Such charts help identify gaps in available content, highlight popular learning areas, and support better decision-making when choosing courses. Data visualization is important because it improves comprehension, reveals hidden patterns, and enhances the overall user experience by turning complex information into meaningful graphical summaries. Below is the snapshot of the data represented in graphical format.

Course Subject Distribution



Difficulty Pie Chart



Recommendation Approach: Content-based filtering + Collaborative filtering

The code combines **content-based filtering** and a simplified **collaborative filtering** approach to generate personalized course recommendations. In the content-based method, the system analyzes each course’s textual information—specifically *tags* and *subject* fields—using TF-IDF vectorization. This transforms the text into numerical vectors that capture how important each word is. The cosine similarity between these vectors is then calculated, allowing the system to identify courses that are most similar to the one selected by the user. This ensures that learners receive recommendations aligned with their known interests.

For collaborative filtering, the code simulates user–course interactions by generating a sample ratings dataset. It identifies which courses the logged-in user has already interacted with and then assigns random prediction scores to all unseen courses. Although simplified, this mimics how real systems use patterns from multiple users to suggest new content. Combining both approaches improves recommendation quality by balancing user preferences with broader learning patterns.

Personalized Recommendations

Select a course you liked:

How to Learn Online

	title	subject	tags
249	Introduction to MachineLearning and AI	Education & Teacher Training	None
313	Introduction toSocial-Emotional Learning(SEL)	Education & Teacher Training	None
325	50-hour English LanguageTeaching Methodology	Education & Teacher Training	None
332	E-COMMERCEFUNDAMENTALS	Education & Teacher Training	None
520	Programming 101: AnIntroduction to Python forEducators	Education & Teacher Training	None

These are the snapshots of recommendations system which is presented in a drop-down menu and in a tabular format.

Collaborative Recommendations

Choose a demo user:

alice

	title	subject	pred_score
338	Blockchain: UnderstandingIts Uses and Implications	Business & Management	0.9997
691	Semiconductors forBeginners	Engineering	0.9969
339	Gesti3n 3gil de proyectos	Business & Management	0.9966
281	Basic Spanish 3: Gettingthere	Language	0.993
194	UX Design	Design	0.9905

Deployment: Build a recommendation dashboard where students log in and get personalized course lists.

To deploy my recommendation system, I created a fully interactive **Streamlit dashboard** that allows students to log in and instantly access personalized course suggestions. The dashboard integrates data processing, user interaction, authentication, recommendation logic, and visualization, all within a clean and responsive interface.

I used **Streamlit** as the main deployment framework because it enables fast, Python-based web app creation without requiring HTML, CSS, or JavaScript. This allowed me to transform the machine learning logic into an accessible, user-friendly application. Streamlit components such as `st.text_input`, `st.selectbox`, `st.write`, `st.pyplot`, and `lalt` elements helped structure the dashboard into clear sections for login, course browsing, recommendations, and analytics.

For authentication, instead of using Streamlit Authenticator (which caused compatibility issues), I implemented a **custom login system** using simple Python dictionaries and session state. This keeps the login system lightweight while still enabling username/password verification and session persistence. The logic checks submitted credentials against predefined usernames and hashed passwords, ensuring only authorized users access the dashboard.

Behind the scenes, I used **pandas** to load and preprocess the dataset, clean column names, and filter courses based on user-entered keywords. I employed **NumPy** for randomization and handling arrays during the recommendation steps.

For the recommendation engine, I combined two methods:

- 1. Content-Based Filtering:**

I applied **TF-IDF vectorization** using `TfidfVectorizer` from `scikit-learn` to convert course descriptions, subjects, and tags into numerical vectors. I then used **cosine similarity** and `linear_kernel` to calculate how similar courses are to one another. This allows the system to recommend courses that closely match the student's interests based on course content.

- 2. Collaborative Filtering (Simplified):**

I generated a demo ratings dataset with simulated user-course interactions.

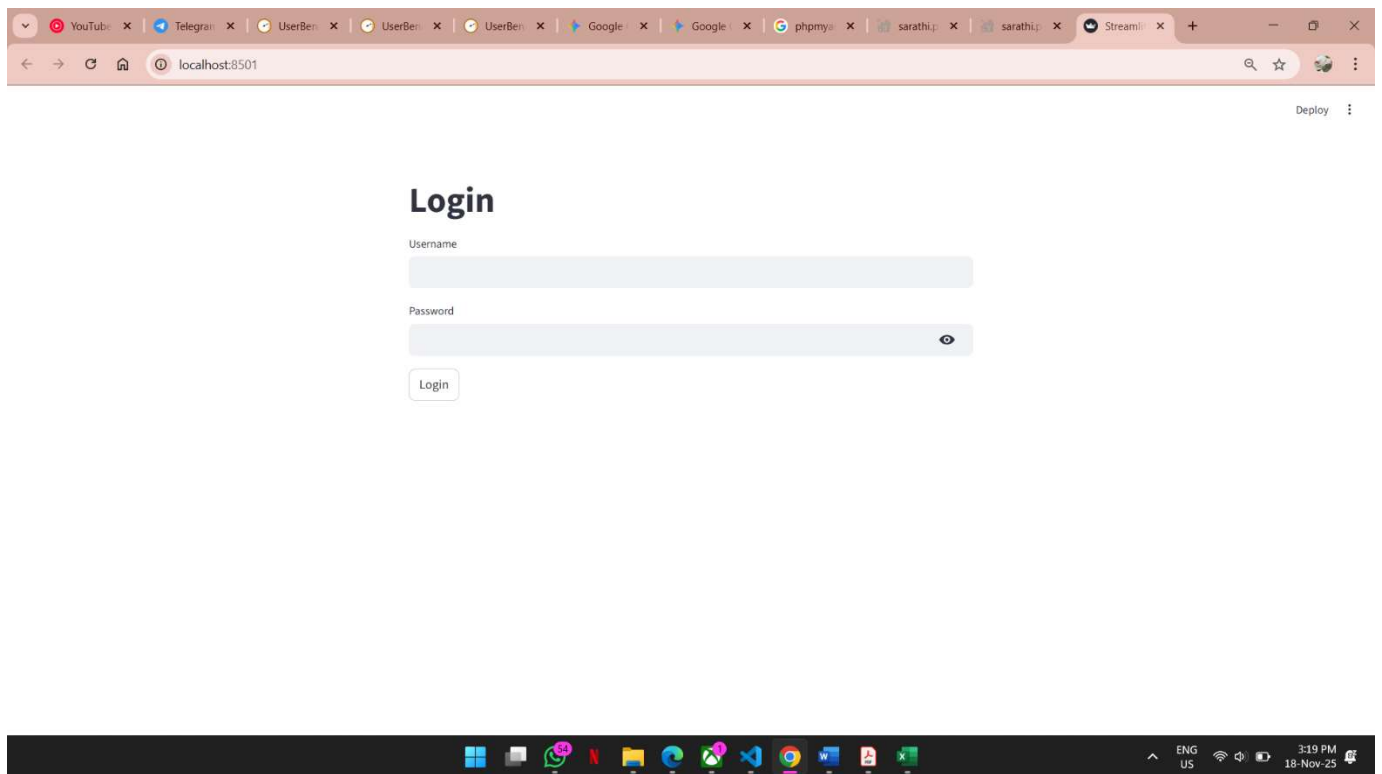
Unseen courses were assigned computed or random prediction scores, mimicking how collaborative filtering algorithms estimate user preferences based on patterns from multiple learners.

To support deployment-friendly analytics, I incorporated visually rich charts using **matplotlib**. Bar graphs and pie charts visualize subject distribution and course difficulty, helping users understand the dataset at a glance.

Once the dashboard logic was complete, I placed it into a single Python file that runs seamlessly as a local or hosted web app using streamlit run. This allows effortless deployment on platforms such as Streamlit Cloud, GitHub Pages (via Streamlit), Render, or other hosting services that support Python applications.

Overall, the deployment integrates data engineering, machine learning, backend login logic, and front-end visualization into one cohesive, accessible, and intelligent course recommendation system.

1. Entering the dashboard you are greeted with the login page



- Upon entering the correct credentials, you enter the dashboard with this view

Course Recommendation System

Browse & Search Courses

Enter preferred subject or tag:

Personalized Recommendations

Select a course you liked:

How to Learn Online

	title	subject	tags
249	Introduction to Machine Learning and AI	Education & Teacher Training	None
313	Introduction to Social-Emotional Learning (SEL)	Education & Teacher Training	None
325	50-hour English Language Teaching Methodology	Education & Teacher Training	None
332	E-COMMERCE FUNDAMENTALS	Education & Teacher Training	None
520	Programming 101: An Introduction to Python for Educators	Education & Teacher Training	None

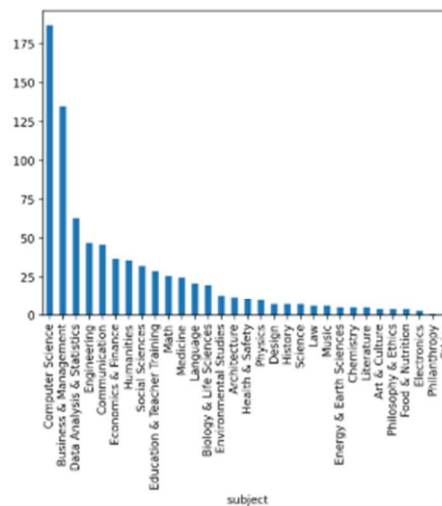
Collaborative Recommendations

Choose a demo user:

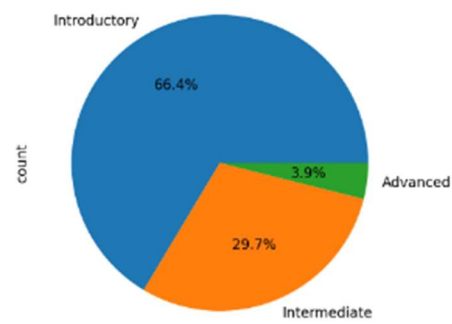
alice

	title	subject	pred_score
338	Blockchain: Understanding its Uses and Implications	Business & Management	0.9997
691	Semiconductors for Beginners	Engineering	0.9969
339	Gestión Ágil de proyectos	Business & Management	0.9966
281	Basic Spanish 3: Getting there	Language	0.993
194	UX Design	Design	0.9905

Course Subject Distribution



Difficulty Pie Chart



3. Now in the box below **Browse & Search Courses** you enter your preferred subject or tag and then you can see this. Here I have entered operations and we can see the list of courses matching our tag in the dataset ie the excel file from Kaggle.

Browse & Search Courses

Enter preferred subject or tag:

operations

	title	subject	difficulty
9	EntrepreneurialOperations: Launching aStartup	Business & Management	Intermediate
18	Data Science and AgileSystems for ProductManagement	Business & Management	Intermediate
36	Entrepreneurship inEmerging Economies	Business & Management	Introductory
79	Excel: Fundamentos yherramientas	Computer Science	Introductory
123	Data Structures &Algorithms I: ArrayLists,LinkedLists, Stacks andâ€¦	Computer Science	Intermediate
141	Introduction to CloudInfrastructureTechnologies	Computer Science	Introductory
168	Data Science forConstruction, Architectureand Engineering	Data Analysis & Statistics	Introductory
191	Data Engineering Basicsfor Everyone	Data Analysis & Statistics	Introductory
229	Supply Chain Planning andStrategy	Business & Management	Introductory
233	Introduction to DevOpsand Site ReliabilityEngineering	Computer Science	Introductory

You also have the option to manually type the course and get the results or use the collaborative recommendation as shown in the snapshot below.

Personalized Recommendations

Select a course you liked:

How to Learn Online

	title	subject	tags
249	Introduction to MachineLearning and AI	Education & Teacher Training	None
313	Introduction toSocial-Emotional Learning(SEL)	Education & Teacher Training	None
325	50-hour English LanguageTeaching Methodology	Education & Teacher Training	None
332	E-COMMERCEFUNDAMENTALS	Education & Teacher Training	None
520	Programming 101: AnIntroduction to Python forEducators	Education & Teacher Training	None

Collaborative Recommendations

Choose a demo user:

alice

	title	subject	pred_score
338	Blockchain: UnderstandingIts Uses and Implications	Business & Management	0.9997
691	Semiconductors forBeginners	Engineering	0.9969
339	GestiÃ³n Ãgil de proyectos	Business & Management	0.9966
281	Basic Spanish 3: Gettingthere	Language	0.993
194	UX Design	Design	0.9905

This is how I created the dashboard for a **Personalized E-Learning Recommendation System** it was a fun, innovative and knowledgeable experience for me and I got to learn many new things and technologies.