ISSN 1693-9220 (Print) | 2987-8713 (Online), Volume 20, No. 2 (2026), pp. 106-119

International Journal of Enterprise Modelling

Published by: International Enterprise Integration Association

Content-Based Filtering Using TF-IDF for a Course Recommendation System for Indonesian MSME Entrepreneurs

Rr Octanty Mulianingtyas[1], Bagus Hendra Saputra[2], Rohani Situmorang[1], Galih Prakoso Rizky A[1]

[1] Faculty of Computer Science, Universitas Pembangunan Nasional Veteran, Jakarta, Indonesia

[2] Faculty of Engineering and Technology Defence, Universitas Pertahanan, Bogor, Indonesia

Abstract

Micro, Small, and Medium Enterprises (MSMEs), particularly those led by female entrepreneurs, play a vital role in Indonesia's economic development; however, digital learning platforms often lack adaptive mechanisms that align course offerings with individual learning needs. Existing platforms generally rely on manual selection or generic categorization, creating a gap in personalized recommendation support within digital entrepreneurship education. The primary objective of this study is to develop and assess a content-based course recommendation system for Femalepreneur.id that addresses this limitation. A quantitative experimental research design was adopted using user profile data collected through questionnaires and course descriptions obtained from the platform repository. The methodology integrates systematic text preprocessing, feature representation using Term Frequency–Inverse Document Frequency (TF-IDF), and similarity computation through cosine similarity to generate personalized recommendations. The experimental results indicate Mean Precision@3 at 51.5%, Mean Average Precision@3 (MAP@3) at 42.9%, and Hit Ratio@3 at 90.9%. The precision matrix demonstrates the system can recommend relevant result until three courses as the maximum value based on the ground truth. While, Hit Ratio matrix reveals that at least the system can recommend at least one relevant topic. These findings confirm the effectiveness of TF-IDF in modelling textual learning features and highlight the contribution of the proposed system in strengthening personalized digital entrepreneurship learning for female entrepreneurs.

Corresponding Author:

Rr Octanty Mulianingtyas,
Fakultas Ilmu Komputer,
Universitas Pembangunan Nasional Veteran, Jakarta,
Jl. RS Fatmawati No. 1, Pondok Labu, Jakarta Selatan12450, Indonesia
rroctantymulianingtyas@upnvj.ac.id

Article Info

Article history:

Received : Feb 09, 2026
Revised : Feb 61, 2026
Accepted : Apr 22, 2026

Keywords:

Content-Based Recommendation;
Course Recommendation System;
Personalized Learning;
Women Entrepreneurs.

Introduction

Micro, Small, and Medium Enterprises (MSMEs) play a central role in Indonesia's economy, contributing substantially to employment generation, income distribution, and economic resilience (Salsabillah et al., 2023). National statistics consistently indicate that MSMEs dominate the business structure and absorb a large share of the workforce (Sinha et al., 2024; Tambunan, 2023). Within this sector, women entrepreneurs represent a rapidly expanding yet structurally vulnerable group, often

facing limited access to continuous and context-relevant entrepreneurial learning. Although digital education platforms help overcome geographical and temporal barriers (Chen et al., 2021; Chotisarn & Phuthong, 2026), their effectiveness depends on the alignment between learning content and users' business stages, competencies, and sectoral needs (Martínez-martínez et al., 2025).

Femalepreneur.id was established to address these challenges by providing online courses, live sessions, business diagnostics, and curated articles tailored to women entrepreneurs in Indonesia. However, as the platform expands, users increasingly encounter information overload, complicating the identification of relevant courses (Alshammary & Alhalafawy, 2023). Despite having structured user profiles and detailed course descriptions, the platform does not yet implement a personalized recommendation mechanism. Consequently, users receive uniform course exposure regardless of entrepreneurial background or learning objectives. Prior research shows that the absence of personalization in digital learning environments is associated with lower engagement and suboptimal learning outcomes (Nguyen et al., 2024).

Recommendation systems have become integral to modern e-learning ecosystems, aiming to map item attributes to user preferences in order to generate relevant recommendation lists (Huang, 2023). Empirical evidence demonstrates that personalized recommendations can enhance engagement, reduce cognitive load, and improve sustained learning outcomes (Sinha et al., 2024). In this domain, Collaborative Filtering and Content-Based Filtering remain dominant approaches, while Hybrid Filtering integrates both (Harun et al., 2025; Saifudin & Widiyaningtyas, 2024). Collaborative Filtering depends heavily on historical interaction data and may underperform in cold-start or sparse-data contexts. In contrast, Content-Based Filtering leverages explicit item attributes, making it more robust in environments characterized by new users and rich textual descriptions (Tambunan, 2023).

Although recommendation techniques have been widely applied in general e-learning and commercial platforms, recent empirical studies predominantly focus on higher education, retail, or generic entrepreneurship users. There remains limited research specifically addressing personalized learning recommendation systems for women-led MSMEs on Indonesian edutech platforms. This gap is particularly critical given Indonesia's status as a developing economy where digital inclusion and women's entrepreneurial empowerment are strategic national priorities within global discussions on inclusive digital entrepreneurship. Existing studies rarely examine how content-based models can be operationalized in platforms serving heterogeneous women entrepreneurs with limited interaction histories.

Given these contextual and empirical gaps, a Content-Based Filtering approach using TF-IDF and cosine similarity is theoretically and practically justified. TF-IDF offers interpretable and computationally efficient textual feature representation (Borbon & Ebardo, 2025; Zhou et al., 2024), and has demonstrated effectiveness in classification and clustering tasks (Guleria, 2025; Wahyuningsih & Chen, 2024; Luo & Lu, 2024). Cosine similarity is widely adopted for measuring similarity in high-dimensional sparse vector spaces within text-based recommendation systems (Rinjeni et al., 2024).

Accordingly, this study seeks to answer the following research question: To what extent a Content-Based Filtering model using TF-IDF and cosine similarity can improve the relevance of course recommendations for women entrepreneurs on Femalepreneur.id? By addressing this question, the study contributes both methodologically through the application of text-based recommendation techniques in a gender-focused MSME context and practically by supporting inclusive digital entrepreneurship development in Indonesia.

Methods

Overview

This study adopts a user-to-item content-based filtering approach to develop a personalized course recommendation system for women entrepreneurs on the Femalepreneur.id platform. The proposed method relies on the semantic similarity between user profiles and course descriptions to identify courses that are most relevant to individual learning needs. The methodological framework is designed to be systematic and reproducible, consisting of data acquisition, preprocessing, feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), similarity computation using Cosine Similarity, recommendation generation, and model evaluation (Precision@K, Mean of Average Precision@K, Hit Rate@K). The overall system architecture and data flow are illustrated in Figure 1, which depicts the relationship between user data, course data, and the recommendation engine.

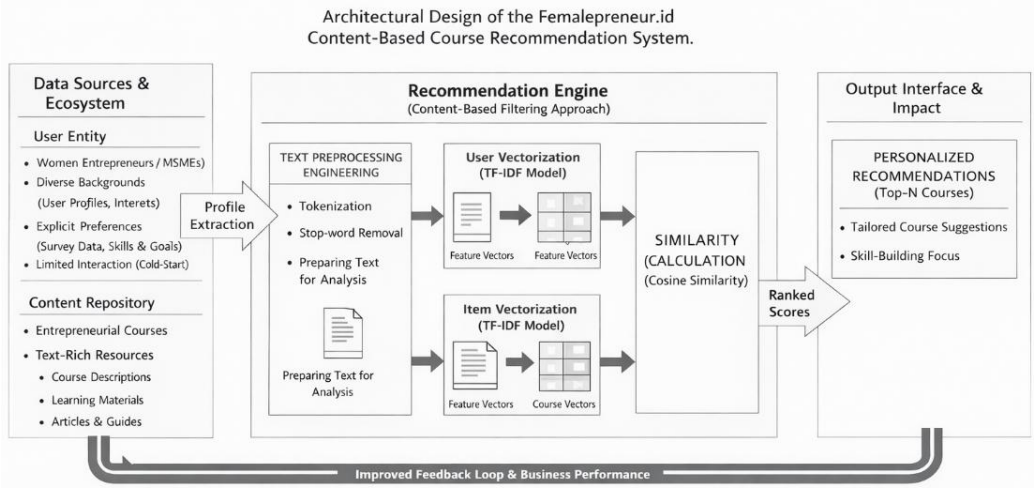

Figure 1. The Recommendation System Architecture

Figure 1 illustrates how the recommendation system works. The recommendation system will receive the input from the user profile, including the length of developing their business, range of age, learning topics, and user budget for buying a course. The system will display personalized recommendation courses from the content repository based on the text vectorization process and item ranking according to the content similarity.

Data Collection

Data were collected directly from active users of the Femalepreneur.id platform through an online questionnaire distributed to registered women entrepreneurs from 2021 to 2025. This questionnaire was designed to capture explicit user preferences and background characteristics relevant to learning needs and business development, ensuring that the recommendation system reflects users' actual expectations rather than relying solely on implicit behavioural data. A total of 98 valid responses were obtained and used in this study. In parallel, course-related data were extracted from the Femalepreneur.id course catalogue, including course description and price, to support the content-based recommendation process.

All collected data were anonymised before processing to preserve user privacy and comply with ethical research standards. Attributes representing direct user identity, such as email addresses, were removed, as they are irrelevant to the recommendation task. A full description of the data set structure is shown in Table 1.

Table 1. The Dataset

LamaBisnis	Harga Course	TipeOnlineCourse	TertarikBelajarBisnis	PrioritasMemilihKelas	TopikPembelajaran	Fitur Diharapkan
Sudah memulai usaha (1-3 tahun)	Rp.500.000 - Rp.1.000.000	Video Rekaman	Yes	Popularitas/ rating, Sertifikat	Kepemimpinan, Pengembangan Diri, Keuangan, Marketing, Branding, Pembiayaan	Pembiayaan dan pemasaran
Memiliki Usaha lebih dari 3 tahun	Rp. 200.000 - 500.000	Kelas Live	Yes	Modul, Mentor, Jumlah Peserta yang Mengikuti, Popularitas/ rating, Sertifikat	Kepemimpinan, Keuangan, Manajemen Bisnis, Pembiayaan, Produk Kerajinan	Tentang pameran pameran produk
Belum memiliki Usaha, dan ingin membangun usaha	Rp. 200.000 - 500.000	Video Rekaman	Yes	Modul, Mentor	Pengembangan Diri, Public Speaking, Keuangan, Manajemen Bisnis	Template siap pakak

Table 1 describes all attributes used in the dataset and the data involved. The dataset combines user profile information and course attributes, enabling direct comparison between user profiles and item profiles within a user-to-item content-based filtering framework. Several user attributes were selected as relevant features for the recommendation process and subjected to data preprocessing. These attributes include: LamaBisnis, representing the length of the user's business experience (no business yet, 1–3 years, or more than 3 years); HargaCourse, reflecting the affordable price range of a course, and TipeOnlineCourse, indicating preference for recorded video or live class formats. Additional attributes include: TertarikBelajarBisnis, representing the user's interest in studying business; PrioritasMemilihKelas, describing decision priorities when selecting courses; TopikPembelajaran, covering preferred learning topics such as marketing, business management, leadership, finance, branding, and digital products; and FiturDiharapkan, representing expected platform features.

The proposed model computes similarity between user profiles and course profiles using a user-to-item similarity approach, rather than user-to-user similarity, allowing the system to generate personalized course recommendations that align with individual learning preferences and business characteristics. The entire steps of the proposed model are represented in Fig. 2.

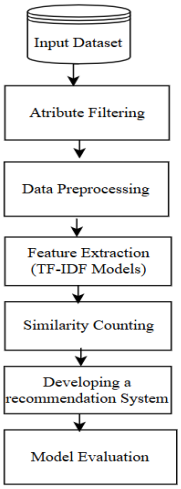

Figure 2. The proposed Model

Fig. 2 illustrates the procedure for constructing the proposed model, which begins with filtering data by selected attributes, pre-processing the data to ensure the dataset's eligibility, feature extraction, developing a recommendation system based on the feature extraction results, and subsequently evaluating the model using several relevant matrices.

Attribute Filtering

The dataset is filtered by two attributes, involving TertarikBelajarBisnis (the user's interest in learning business) and HargaCourse (the user's capability for buying a course). For TertarikBelajarBisnis, the dataset is filtered only for users who express interest in learning business. For HargaCourse. The filtering process decreases the amount of data from 98 to 95.

Data Pre-processing

Data pre-processing was conducted to ensure data consistency, reduce noise, and improve the quality of textual feature representation used in the recommendation process. Textual preprocessing was applied primarily to attributes containing free-text input, particularly FiturDiharapkan, as this attribute directly reflects user-generated textual preferences. The first step was case folding, in which all text was converted to lowercase to ensure uniform text representation and prevent inconsistent treatment of identical words with different capitalisation. Next, noise removal was performed to eliminate punctuation marks, special characters, URLs, emoticons, excessive whitespace, and numerical tokens that do not contribute semantic value to the recommendation task. Following noise removal, stopword removal was applied using an Indonesian stopword list provided by the NLTK library, as the dataset is predominantly in the Indonesian language. This process removes frequently occurring function words that carry limited discriminative meaning, thereby improving the signal-to-noise ratio in text representation. The cleaned text was then subjected to tokenization, where sentences were segmented into individual terms or tokens, forming the basis for feature extraction. Finally, stemming was conducted using an Indonesian stemming algorithm to reduce words to their root forms. This step consolidates morphological variations of words and improves term matching between user profiles and course descriptions.

Feature Extraction Using TF-IDF

In this study, TF-IDF is employed to measure the importance of a term within a document relative to the entire corpus. TF-IDF is a form of two measurements, i.e., Term Frequency and Inverse Document Frequency. Term Frequency (TF) reflects the occurrence count of a word within a

document. In contrast, Inverse Document Frequency (IDF) evaluates how informative a word is by comparing the total number of documents to the number of documents in which the word appears, resulting in smaller IDF values for commonly occurring terms (Jin et al., 2024; Liang & Niu, 2022).

Eq(1) illustrates TF and IDF, where TF refers to term frequency. The relative frequency of a word in a document is measured by comparing its occurrence to the total number of words in that document, as described in Eq (2). The notation $n_{x,y}$ indicates the frequency with which a word occurs in a specific document. In the equation below, inverse document frequency (IDF) is applied to assign higher weights to terms that occur less frequently across all documents in the dataset, as mathematically expressed in Eq (3). According to Eq(4), the TF-IDF computation can be refined, where 'D' corresponds to the total number of documents, $tf_{x,y}$ denotes the occurrence frequency of term x in document y, and df_x indicates the document frequency of term x (Guleria, 2025).

$$tf_{idf} = tf_{x,y} \times idf\ (\omega) \quad (1)$$

$$tf_{x,y} = \frac{n_{x,y}}{\Sigma_{i \in y}\, n_x(i)} \quad (2)$$

$$tf_{x,y} = \frac{n_{x,y}}{\Sigma_{i \in y}\, n_x(i)} \quad (3)$$

$$\omega_{x,y} = tf_{x,y} \times \log \frac{D}{df_x} \quad (4)$$

Source: (Guleria, 2025)

By representing user profiles and course descriptions as weighted vectors in a shared feature space, TF-IDF facilitates effective comparison between textual entities. Recent studies demonstrate that TF-IDF remains a robust and interpretable feature extraction method in content-based recommendation systems, particularly in domains with rich textual descriptions (Hiro et al., 2023).

Similarity Counting

In the context of this study, Cosine Similarity quantifies the semantic alignment between a user's learning preferences and a course's content description. Higher similarity values indicate greater relevance, while lower values indicate limited correspondence. The use of Cosine Similarity quantifies semantic alignment between user learning preferences and course descriptions, where higher similarity scores indicate stronger relevance. Its computational efficiency and robustness in sparse data environments are well documented in recent recommendation system research (Raharjo & Arifin, 2023).

Developing a Recommendation System

For a given user, courses are ranked based on their similarity scores in descending order. The top-ranked courses are then selected as personalised recommendations. This user-to-item matching process ensures that recommended courses align closely with individual user characteristics, including business experience, user preferences, range budget for one course, and preferred learning topics. The recommendation output is designed to be interpretable, enabling platform administrators to understand the basis of each recommendation and facilitating future system refinement.

Evaluation of Model

Model performance was evaluated using standard metrics, namely Precision@k, Mean Precision@k Average Precision@k, Mean of Average Precision@k, dan Hit Rate@k. Accuracy is not utilized in this study because the proposed content-based filtering approach does not perform classification tasks. The evaluation procedure was designed to ensure methodological consistency and reproducibility, without introducing result interpretation or comparative analysis at this stage.

a. Precision@k

Precision evaluates the performance of a recommendation system by determining the ratio of relevant items correctly recommended to the total number of items suggested, where non-relevant items are considered false positives (Ertuğrul & Bitirim, 2025). The corresponding formula is depicted in Eq. (5).

$$Precision@k = \frac{(\text{\# of recommended items @k that are relevant})}{\text{\# of recommend items @k}} \quad (5)$$

Source : (Ertuğrul & Bitirim, 2025)

b. Mean Precision@k

Mean Precision@k represents the average of Precision@k divided by total number of users. The corresponding formula is depicted in Eq. (6)

$$Mean\ Precision@k = \frac{\text{Total number of Precision@k}}{\text{Total number of users}} \quad (6)$$

Source : (Ertuğrul & Bitirim, 2025)

d. Average Precision@k (AP@k)

It measures how effectively the system orders relevant items within its recommendations by taking into account the precision at every position where a relevant item appears in the ranked list. In this formulation, k indicates the rank position of the retrieved items, while n represents the total number of selected items. P(k) refers to the precision value at rank k. Furthermore, rel(k) is an indicator function that equals 1 when the item at position k is relevant and 0 otherwise, and Nr denotes the total count of relevant items. The corresponding formula is depicted in Eq. (7).

$$Average\ Precision@k = \frac{\sum_{k-1}^{n} P(k).rel(k)}{n_r} \quad (7)$$

Source : (Ertuğrul & Bitirim, 2025)

e. *Mean Average Precision (MAP)*

MAP represents the mean of precision values computed for each user and then averaged across all users. When there are K recommended items, the precision P(k) is calculated and averaged over all positions from 1 to K. The related formulation is shown in Equation (8).

$$MAP@k = \frac{\sum_{i-1}^{I} AP_i}{I} \ (9) \quad (8)$$

Source : (Ertuğrul & Bitirim, 2025)

f. Hit Ratio@k (Hit R@k)

Hit Ratio (HitR) in a recommender system refers to the proportion of users for whom the relevant item(s) appear within a recommendation list of length L. As the value of L increases, the Hit Ratio also tends to rise, since the likelihood of including the correct item(s) in the recommendation list becomes greater. Therefore, selecting an appropriate value of L is a crucial step in evaluating the performance of a recommender system (Ertuğrul & Bitirim, 2025).

$$Hit\ Ratio\ @k = \frac{|\ \cup_{hit}^{L}\ |}{\cup\ all} \quad (10)$$

Results and Discussion

Data Pre-processing and Dataset Refinement

This section evaluates the impact of data pre-processing on dataset quality and modelling readiness. Pre-processing is essential in text-based recommendation systems because raw user-generated data often contains inconsistencies and noise that can distort feature representation. The initial dataset comprised 98 user records collected through online questionnaires on Femalepreneur.id, which underwent filtering and cleaning prior to feature extraction. Following

filtering, the dataset was reduced to 95 records, primarily due to attribute selection based on TertarikBelajarBisnis and HargaCourse. Although the dataset size decreased, this refinement improved data reliability by retaining more informative and consistent records.

Results of Textual Data Pre-Processing

This attribute captures users' expectations regarding course characteristics and platform functionality, making it a key semantic indicator in the recommendation process. Fig. 3 illustrates the outcome of the pre-processing pipeline applied to this attribute, including case folding, noise removal, stopwords elimination, tokenization, and stemming.

Fig 3. Result of Data Preprocessing in the attribute of FiturDiharapkan

As shown in Figure 3, the pre-processing steps successfully transformed heterogeneous and unstructured textual inputs into normalized and linguistically consistent tokens, particularly for the attribute User's Expected Features (FiturDiharapkan). By reducing lexical variation and eliminating irrelevant tokens, the model becomes more sensitive to genuinely discriminative terms that reflect user preferences. Ultimately, this step improves semantic alignment between user profiles and course descriptions, thereby strengthening the relevance of the recommendation output.

Construction of Text-Based Features

Following pre-processing, composite textual features were constructed to represent user preferences and course content within a shared semantic space. Three attributes were integrated: User's Expected Features, Learning Topics, and User Priority in course selection, as they provide linguistically rich and contextually relevant information reflecting learning needs and decision-making behaviour.

Fig. 4 demonstrates the result of text feature construction from these attributes. The integration of these attributes produces a more comprehensive representation of user intent, where Expected Features capture experiential requirements, Learning Topics represent domain interests, and Priority reflects learning goal importance. This multi-attribute approach supports more meaningful similarity computation and improves recommendation relevance, consistent with findings in e-learning recommendation research (Huang, 2023).

```
df['clean_input']

0     forum diskusi apps topik ganti tambah tahu ala...
1     live chat public speaking, marketing, pengemba...
2     femalepreneur learning core rate paham materi ...
3     materi modul akses video rekam presents live k...
4     apikasi bisnis tingkat pengembangan diri, keua...
                            ...
90    tema tarik branding, pengembangan produk modul...
91    feedback fitur isi lapor monitoring ajar serta...
92     pengembangan diri, keuangan, marketing, manaj...
93    interaksi mentor kepemimpinan, pengembangan di...
94           modal usaha pengembangan produk sertifikat
Name: clean_input, Length: 95, dtype: object
```

Fig 4. Result of building text feature

Fig 4 illustrates the result of the text feature that will be utilized for the text vectorizer process by the TF-IDF approach, as it is saved by the tfidf_matrix variable. Several parameters are set for the text vectorizer process, as represented in Fig 5, including ngram_range = (1,2), which indicates that both unigrams and bigrams are processed through textual feature representation. This configuration enables the model to capture individual keywords and meaningful multi-word expressions, thereby improving semantic representation and similarity measurement in text-based recommendation systems. L2 norm to prevent the biased result from the document length. The word pattern will be a parameter for the measurement of TF IDF, instead of the document length. Then, it will make the cosine similarity measure consistent and valid.

TF-IDF Vectorisation and Cosine Similarity Analysis

Following feature construction, TF-IDF vectorization was applied to transform textual representations into numerical vectors. This stage aims to quantify the importance of terms within user profiles and course descriptions by assigning higher weights to distinctive semantic elements while suppressing commonly occurring, less informative words. The resulting TF-IDF vectors are inherently high-dimensional and sparse, which makes cosine similarity an optimal choice for similarity measurement (Raharjo & Arifin, 2023).

Higher cosine similarity scores indicate a stronger alignment between a user's learning trajectory and a course's content profile, providing a stable and interpretable foundation for content-based recommendation. These observed similarity patterns, which are further illustrated in Fig. 5, align with recent findings in recommendation system research regarding the robustness of hybrid textual feature representation.

```
from sklearn.feature_extraction.text import TfidfVectorizer

tfidf = TfidfVectorizer(
    ngram_range=(1,2),
    min_df=1,
    max_df=0.9,
    sublinear_tf=True,
    norm='l2',
    max_features=5000
)

tfidf_matrix = tfidf.fit_transform(df['clean_input'])
```

Fig 5. Text Vectorization process

Fig. 5 illustrates the transformation pipeline that converts raw textual attributes from user profiles and course descriptions into TF-IDF-weighted vectors through preprocessing and vectorisation, producing structured numerical representations suitable for similarity computation. This approach strengthens term discrimination by increasing the weight of distinctive learning preference indicators while reducing the influence of common terms, consistent with previous findings on TF-IDF effectiveness. However, because TF-IDF relies on lexical similarity, it may not

fully capture deeper contextual meaning, potentially limiting adaptability for implicit learning preferences (Hiro et al., 2023). The subsequent step involves calculating similarity using cosine similarity, with the results presented in Fig. 6.

```
array([[1.        , 0.07769569, 0.05183048, ..., 0.2129847 , 0.19662539,
        0.01917287],
       [0.07769569, 1.        , 0.10414107, ..., 0.0898732 , 0.10690606,
        0.0460284 ],
       [0.05183048, 0.10414107, 1.        , ..., 0.08214638, 0.0448736 ,
        0.01023419],
       ...,
       [0.2129847 , 0.0898732 , 0.08214638, ..., 1.        , 0.18360183,
        0.06531992],
       [0.19662539, 0.10690606, 0.0448736 , ..., 0.18360183, 1.        ,
        0.03351873],
       [0.01917287, 0.0460284 , 0.01023419, ..., 0.06531992, 0.03351873,
        1.        ]])
```

Fig 6. The result of counting similarity

Fig. 6 presents the ranked cosine similarity values generated by comparing a selected course against other course items in the dataset. The figure illustrates how the recommendation engine orders courses based on similarity distance, where higher similarity scores indicate closer semantic correspondence. The ranking mechanism enables prioritisation of courses that closely match user learning preferences and topical interests. From a practical standpoint, this supports Femalepreneur.id users by presenting learning options that are most relevant to their entrepreneurial development needs. Nevertheless, similarity-based ranking may be sensitive to overlapping textual descriptions across multiple courses, potentially reducing thematic diversity in recommendation outputs.

Fig. 7 presents the final recommendation results generated after integrating TF-IDF similarity computation with attribute-based filtering. The figure demonstrates how personalised course recommendations are produced based on combined semantic relevance and profile compatibility. This integration reflects established content-based recommendation practices that combine textual similarity with contextual user attributes to improve recommendation precision.

```
result = recommend_topics_with_similarity(
    df=df,
    cosine_sim=cosine_sim,
    user_lama_bisnis= "Sudah memulai usaha (1-3 Tahun)",
    user_price= "Rp. 200.000 - Rp 500.000",
    user_topics= ["Kepemimpinan", "Manajemen Bisnis", "Branding"]
)
result
```

```
Total Result: 43
{'LamaBisnis': 'Sudah memulai usaha (1-3 Tahun)',
 'HargaCourse': 'Rp. 200.000 - Rp 500.000',
 'TopCourseSimilarity': [{'Index': 45,
   'LamaBisnis': 'Sudah memulai usaha (1-3 tahun)',
   'HargaCourse': 'Tidak Berkenan Membayar',
   'Topics': 'pengembangan diri, public speaking, keuangan, marketing, branding, legal, manajemen bisnis, pembiayaan, pengembangan produk',
   'TopicCount': 9,
   'Similarity': np.float64(0.276)},
  {'Index': 69,
   'LamaBisnis': 'Sudah memulai usaha (1-3 tahun)',
   'HargaCourse': 'Rp. 0 - Rp. 200.000',
   'Topics': 'public speaking, keuangan, marketing, branding, manajemen bisnis, pengembangan produk',
   'TopicCount': 6,
   'Similarity': np.float64(0.227)},
  {'Index': 35,
   'LamaBisnis': 'Sudah memulai usaha (1-3 tahun)',
   'HargaCourse': 'Rp. 0 - Rp. 200.000',
   'Topics': 'pengembangan diri, public speaking, marketing, manajemen bisnis, pengembangan produk',
   'TopicCount': 5,
   'Similarity': np.float64(0.2173)},
```

Fig 7. Result of Recommendation System for one of user's backgrounds

As shown in Fig. 7, from the application perspective, the system supports Femalepreneur.id users in identifying learning resources aligned with business maturity, financial

capacity, and professional development goals. However, the effectiveness of the final recommendations remains dependent on the completeness and quality of user profile information and course descriptions, as limited or inconsistent textual data may reduce recommendation accuracy (Raharjo & Arifin, 2023).

Evaluation Results and Metric Interpretation

The recommendation model was evaluated using Precision, Recall, and F1-score, which provide reliable performance assessment for a Recommendation System (Chen et al., 2021) (Ertuğrul & Bitirim, 2025). The model achieved balanced performance with Mean Precision@3 at 51.5%, Mean Average Precision@3 (MAP@3) at 42.9%, and Hit Rat@3 at 90.9% that represents in Fig. 10.

The high Hit Ratio indicates strong alignment between recommended courses and user preferences, while the Mean confirms adequate retrieval of relevant learning materials. The balance of these metrics demonstrates the system's effectiveness in maintaining recommendation relevance and coverage, consistent with evaluation standards in educational recommender systems (Kabir et al., 2023).

Prior studies in educational and content-based recommender systems consistently emphasise the superiority of Precision and F1-score over accuracy when relevance quality and user satisfaction are prioritised (Li et al., 2021).

Methodologically, the results confirm that TF-IDF effectively captures discriminative textual features from user profiles and course descriptions, while cosine similarity provides stable similarity measurement in high-dimensional sparse spaces. The training performance shown in Fig 11 further indicates consistent metric distribution, supporting the robustness of similarity-based textual modelling for personalised recommendations.

Precision@K is used as a metric because we can measure the number of Precision of model by value of K. The precision value is obtained from the number of relevant recommended topics compared to number of recommended topics that is represented. We selected 3 as a value of K since the recommendation result display top-3 value that adjust the total number of the ground truth data for each user's background that comprises 3 topics. For example, the user who has characteristics: has developed business for 1-3 years, allocating budget for buying course in the range of Rp. 200.000 -Rp. 500.000. The evaluation result as shown in Fig. 10 suggests that personalized recommendations can reduce information overload and enhance user engagement, although performance remains dependent on the completeness of user-provided textual preferences (Chen et al., 2021; Kabir et al., 2023).

LamaBisnis	HargaCourse	TotalResults	Predicted	Ground Truth	Number of Relevance	Hit@3	Precision@3	Average_Precision@3
Belum memiliki Usaha, dan ingin membangun usaha	Tidak Berkenan Membayar	[17]	pengembangan diri, keuangan, marketing	pengembangan diri, pengembangan produk, keuangan	2	1	0.667	0.666667
Belum memiliki Usaha, dan ingin membangun usaha	Rp. 0 - Rp. 200.000	[17]	pengembangan diri, manajemen bisnis, keuangan	keuangan, legal, kepemimpinan	1	1	0.333	0.111111
Belum memiliki Usaha, dan ingin membangun usaha	Rp. 200.000 - Rp. 500.000	[15]	pengembangan diri, manajemen bisnis, keuangan	kepemimpinan, pengembangan diri, manajemen bisnis	2	1	0.667	0.666667
Sudah memulai usaha (1-3 tahun)	Tidak Berkenan Membayar	[47]	pengembangan diri, manajemen bisnis, marketing	keuangan, kepemimpinan, pengembangan produk	0	0	0.000	0.000000
Sudah memulai usaha (1-3 tahun)	Rp. 0 - Rp. 200.000	[31]	pengembangan diri, manajemen bisnis, marketing	keuangan, kepemimpinan, pengembangan diri	1	1	0.333	0.333333
Sudah memulai usaha (1-3 tahun)	Rp. 200.000 - Rp. 500.000	[43]	pengembangan diri, manajemen bisnis, marketing	kepemimpinan, manajemen bisnis, branding	1	1	0.333	0.166667
Sudah memulai usaha (1-3 tahun)	Rp. 500.000 - Rp. 1.000.000	[47]	pengembangan diri, manajemen bisnis, marketing	manajemen bisnis, marketing, public speaking	2	1	0.667	0.388889
Memiliki Usaha lebih dari 3 tahun	Tidak Berkenan Membayar	[23]	pengembangan diri, marketing, pengembangan produk	pengembangan diri, marketing, pengembangan produk	3	1	1.000	1.000000
Memiliki Usaha lebih dari 3 tahun	Rp. 0 - Rp. 200.000	[16]	pengembangan diri, marketing, pengembangan produk	keuangan, kepemimpinan, pengembangan diri	1	1	0.333	0.333333
Memiliki Usaha lebih dari 3 tahun	Rp. 200.000 - Rp. 500.000	[23]	pengembangan diri, marketing, pengembangan produk	pengembangan diri, marketing, pengembangan produk	3	1	1.000	1.000000
Memiliki Usaha lebih dari 3 tahun	Rp. 500.000 - Rp. 1.000.000	[20]	pengembangan diri, marketing, pengembangan produk	manajemen bisnis, marketing, public speaking	1	1	0.333	0.166667

Fig 10. Result of Evaluation

```
print("Average Precision@3:", evaluasi["Precision@3"].mean())
print("Mean of Average Precision@3:", evaluasi["Average_Precision@3"].mean())
print("Average Hit Rate:", evaluasi["Hit@3"].mean())
```

```
Average Precision@3: 0.5150909090909092
Mean of Average Precision@3: 0.4393939393939394
Average Hit Rate: 0.9090909090909091
```

Fig 11. Result of Each Metric Evaluation

The Mean Precision@3 value is 0.51, indicating that 51% of the recommended topics are relevant to the user's stated preferences. In this evaluation, the system generated 3 predicted topics (Top-3 recommendations). The user provided 3 relevant topics as ground truth. From the 3 predicted topics, the system is proven can match the user's input preferences from 0 to 3 topics. The Average Hit Ratio@3 value is 0.91, indicating almost entire result of the recommendation system can match at least one relevant topic based on ground truth. From the total of recommended result, only one user background that result zero relevant topic.

This result suggests that the recommendation model is able to correctly capture most of the user's interests. However, one recommended topic did not overlap with the user's stated preferences. This may occur because:

i) The similarity-based ranking prioritizes frequently occurring topics in highly similar courses.
ii) Some topics are semantically related but not lexically identical (e.g., "Digital Marketing" vs "Marketing").
iii) The frequency-based aggregation of top course topics may introduce slightly broader themes.

Overall, the result of all evaluation matrices indicates that the model demonstrates good relevance performance, although there is still room for improvement in aligning recommendations more precisely with explicit user preferences.

Critical Discussion of Model Performance

The evaluation indicates stable model performance without significant overfitting or underfitting. TF-IDF effectively captures discriminative textual features while avoiding excessive dependence on rare lexical patterns, enabling consistent similarity differentiation between user and course profiles. The use of explicit user attributes also improves interpretability and transparency, providing clearer justification for recommendation results compared with collaborative filtering approaches, which often operate as black-box models. Such transparency is essential in educational decision-support systems and has been widely emphasised in personalised learning research (Kabir et al., 2023).

However, several limitations remain. Reliance on self-reported questionnaire data introduces variability in textual expression that may influence similarity computation, while the relatively small dataset size may restrict generalisability across broader entrepreneurial learning populations, reflecting common constraints in niche educational platforms (Hanaysha, 2022). Despite these limitations, the results confirm that TF-IDF and cosine similarity provide an effective content-based recommendation mechanism for digital learning environments with limited interaction data and diverse user backgrounds.

Conclusion

The evaluation results show that the proposed recommendation model effectively captures user preferences and provides relevant learning topics, with most recommendations aligning well with users' needs and consistently including at least one relevant topic. This supports the proposed personalized framework for Femalepreneur.id, which applies Content-Based Filtering using TF-IDF and cosine similarity to generate context-aware recommendations, particularly in environments with limited interaction data and diverse user backgrounds. By emphasizing explicit preference modeling for women-led MSMEs, the model improves interpretability and helps address cold-start issues. However, some mismatches still occur due to similarity ranking, semantic differences, and topic aggregation, indicating that future improvements should explore hybrid or semantic-based approaches with richer and real-time data to further enhance recommendation accuracy.

Reference

Chen, L., Ifenthaler, D., Yin, J., & Yau, K. (2021). Online and blended entrepreneurship education : a systematic review of applied educational technologies. In *Entrepreneurship Education* (Vol. 4, Issue 2). Springer Singapore. https://doi.org/10.1007/s41959-021-00047-7

Ertuğrul, D. Ç., & Bitirim, S. (2025). Job recommender systems : a systematic literature review , applications , open issues , and challenges. In *Journal of Big Data*. Springer International Publishing. https://doi.org/10.1186/s40537-025-01173-y

Guleria, P. (2025). NLP based text classification using TF-IDF enabled fine-tuned long short-term memory : An empirical analysis. *Array*, *27*(August), 100467. https://doi.org/10.1016/j.array.2025.100467

Hanaysha, J. R. (2022). International Journal of Information Management Data Insights Impact of social media marketing features on consumer ' s purchase decision in the fast-food industry : Brand trust as a mediator. *International Journal of Information Management Data Insights*, *2*(2), 100102. https://doi.org/10.1016/j.jjimei.2022.100102

Hiro, A., Permana, J., & Wibowo, A. T. (2023). *Movie Recommendation System Based on Synopsis Using Content-Based Filtering with TF-IDF and Cosine Similarity*. *9*(2), 1–14. https://doi.org/10.21108/ijoict.v9i2.747

Huang, R. (2023). Improved content recommendation algorithm integrating semantic information. *Journal of Big Data*. https://doi.org/10.1186/s40537-023-00776-7

Jin, Z., Ye, F., Nedjah, N., & Zhang, X. (2024). *Electronic Commerce Research and Applications A comparative study*

of various recommendation algorithms based on E-commerce big data. *68*(October).

Kabir, S., Farrokhvar, L., & Dabouei, A. (2023). A Weakly supervised approach for thoracic diseases detection. *Expert Systems With Applications*, *213*(PB), 118942. https://doi.org/10.1016/j.eswa.2022.118942

Li, H., Wang, Y., Li, Y., Xiao, G., Hu, P., Zhao, R., & Li, B. (2021). *Learning Adaptive Criteria Weights for Active Semi-Supervised Learning*.

Liang, M., & Niu, T. (2022). ScienceDirect ScienceDirect Research on Text Classification Techniques Based on Improved TF- Research on Text Classification Techniques Based on Improved TF- IDF Algorithm and LSTM Inputs IDF Algorithm and LSTM Inputs. *Procedia Computer Science*, *208*, 460–470. https://doi.org/10.1016/j.procs.2022.10.064

Raharjo, M. M., & Arifin, F. (2023). *Machine Learning System Implementation of Education Podcast Recommendations on Spotify Applications Using Content-Based Filtering and TF-IDF*. *8*(November), 221–230.