Building Explainable Recommender System for Engineering Students' Work-Integrated Learning (WIL)

Woorin Hwang
University of Florida
Gainesville, FL, USA
hwangwoorin@ufl.edu

Hongmin (Chip) Li
University of Florida
Gainesville, FL, USA
hli3@ufl.edu

Pauline Aguinalde
University of Florida
Gainesville, FL, USA
aaguinalde@ufl.edu

Yilin Zhang
University of Florida
Gainesville, FL, USA
yilin.zhang@ufl.edu

Jinnie Shin
University of Florida
Gainesville, FL, USA
jinnie.shin@coe.ufl.edu

Kent J. Crippen
University of Florida
Gainesville, FL, USA
kcrippen@coe.ufl.edu

Bruce F. Carroll
University of Florida
Gainesville, FL, USA
bfc@ufl.edu

ABSTRACT

Work-integrated learning (WIL) is essential in engineering education, bridging the gap between theory and practice while enhancing students' employability. This study presents an LLM-enhanced, AI-driven hybrid content-based recommender system designed to optimize WIL experiences by tailoring recommendations based on students' academic and personal backgrounds, as well as the WIL tasks. A sample of 223 undergraduate engineering students participated, providing insights into their WIL experiences. Preliminary results show the internal performance of the recommender system with promising accuracy, and the system is further evaluated to understand its explainability and relevance.

Keywords

Work-integrated learning, LLMs, Recommender System

1. INTRODUCTION

Work-integrated learning (WIL) is a cornerstone of modern engineering education, bridging academic knowledge with real-world applications [2]. Through internships, co-ops, and experiential learning, WIL enhances employability and fosters essential professional skills [5]. It also strengthens students' commitment to the profession by shaping their identity, fostering belonging, and improving self-efficacy through skills development [6]. As engineering education evolves alongside technological advancements, there is a growing need for personalized, data-driven support in WIL settings. However, aligning WIL with academic training and career goals remains a challenge. Many struggle to contextualize
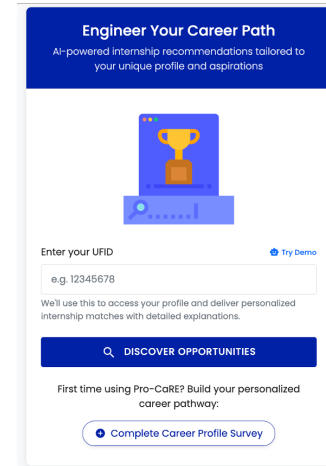
Figure 1: Interface of the WIL recommender system

internship tasks within broader learning objectives, making it difficult to connect their experiences to future roles [16]. Also, the dynamic nature of engineering workplaces complicates the ways to address students' diverse skill development needs and preparing them for emerging industry demands.

Artificial Intelligence (AI)-based recommender systems have emerged as a promising solution in education, transforming how students discover and engage with learning opportunities [15]. These systems leverage machine learning, natural language processing (NLP), and collaborative filtering to analyze learner behaviors, preferences, and performance, generating tailored recommendations that optimize educational experiences [8], such as dynamically adapting learning pathways by recommending courses, reading materials, and assessments to bridge knowledge gaps [3, 12, 11]. By applying similar AI-driven strategies to WIL, there is potential to revolutionize career advising and internship matching, ensuring that students receive personalized guidance tailored

to their skills, interests, and professional goals. Despite their potential in enhancing WIL advising, recommender systems for decision-making remain largely unexplored [10]. Recognizing this gap, this study introduces a scalable AI-based recommender system, the Pro-CaRE Recommender System (Figure 1), to support undergraduate engineering students in making informed career decisions. This study is guided by the following three research questions:

RQ1. How can the AI-based recommender system be designed to align engineering students' technical and professional competencies with WIL opportunities?

RQ2. How accurately can the AI-based recommender system retrieve WIL opportunities where students previously demonstrated high learning growth?

RQ3. How do students perceive the explainability of AI-driven career recommendations, and how might it affect their trust and engagement with the advising process?

2. WIL IN ENGINEERING EDUCATION

Work-integrated learning (WIL) bridges academic coursework with real-world engineering experience through co-ops, internships, and apprenticeships [13]. WIL structures vary based on institutional and industry partnerships, with some programs offering structured objectives while others focus on organic skill development [4]. WIL involves three key stakeholders: students, educational institutions, and employers. Engineering students typically engage in WIL after completing foundational coursework, applying theoretical knowledge in industry settings [9]. Workplace supervisors provide task assignments, professional guidance, and mentoring, influencing students' career trajectories [1]. Educational institutions facilitate WIL through industry collaborations, structured assessments, and career support services. Despite its benefits, WIL programs face challenges in providing consistent student support due to variations in structures, workplace expectations, and mentorship quality. The lack of standardized tools further complicates efforts to ensure relevant learning experiences. Addressing these issues requires data-driven solutions to systematically support students in selecting, preparing for, and maximizing opportunities [10].

3. RECOMMENDATION SYSTEMS

In this section, we introduce the fundamental architecture of our recommender system.

3.1 Content-based Filtering (CBF)

In a content-based recommendation system, *items* (e.g., WIL opportunities) are represented by a set of features, and recommendations are made based on the similarity between items and the student's profile, where each item i is represented as a feature vector $x_i = (x_{i1}, x_{i2}, ..., x_{id}) \in \mathbb{R}^d$. The features in our system include multiple layers of input, such as relevant courses, work settings, company size, and required technical and professional skills. Similarly, the *student* profile, u, is constructed by aggregating the feature vectors of previously interacted items. A common method is the weighted average: $u = \frac{1}{n} \sum_{j=1}^{N} w_j x_j$. The weights are determined based on students' "ratings" of WIL opportunities. In this study, we define these ratings across several dimensions: students' learning outcomes (i.e., technical and professional) and improvements in the affective dimension (i.e., self-efficacy as engineers).

3.2 Collaborative Filtering (CF) Algorithm

To improve recommendation performance, collaborative filtering (CF;[14]) elements are incorporated, which capture user preferences based on the behavior of similar users. Let $R \in \mathbb{R}^{m \times n}$ be the student-item interaction matrix, where $R_{u,i}$ is the explicit learning outcomes of the student u on WIL opportunity (item), i. The new objective is to predict missing values in R, the two models (CBF and CF) can be combined via matrix factorization. In this setting, the student-item rating matrix is decomposed as $R \approx UV^T$, where $U \in \mathbb{R}^{m \times k}$ and $V \in \mathbb{R}^{N \times k}$ represent student and item latent factors, respectively. The item features are incorporated as additional regularization terms, where λ controls for the influence of content features in item representation.

$$min_{u,v} \sum_{(u,i) \in K} (R_{u,i} - U_u V_i^T)^2 + \lambda ||x_i - V_i||^2 \qquad (1)$$

3.3 Enhancement with LLMs

One of the fundamental challenges in designing an effective recommender system for WIL opportunities is the dynamic nature of available opportunities. Unlike traditional course recommendation systems, where the pool of items (courses) remains relatively stable over time, WIL opportunities are constantly evolving—new internships and co-ops are introduced frequently, while others expire. This rapid turnover complicates the process of maintaining an up-to-date and reliable recommendation database. To address this challenge, we leveraged generative Large Language Models—specifically Llama 2—to automatically generate item features for newly introduced WIL opportunities. Instead of relying solely on pre-existing user interaction data, which may be sparse or nonexistent for new items, we employed LLMs to extract and structure core skills and tasks from WIL descriptions, classifying them into 15 categories (Table 2). Given a new WIL description d_i, the LLM generates item features as latent profiles, denoted as: $x_i = LLM(d_i)$. With LLM-enhanced feature generation, we can integrate new opportunities into recommendations immediately, even before any user interaction occurs. The system computes the similarity between a user profile u and newly generated item vectors x_i by evaluating their similarity to other previously experienced and "rated" items based on learning outcomes from existing users. The two feature vectors are then compared to assess their similarity. To provide recommendations, we compute the similarity between a user's profile and a new job posting i using cosine similarity, where the similarity score is determined by the dot product between the user's profile and the new job posting vector.

$$sim(u, x) = \frac{u \cdot x_i}{||u|| \cdot ||x_i||} \qquad (2)$$

3.4 Participants and Data Collection

From March to November 2024, survey questionnaires were distributed via the university's Qualtrics server, yielding a total of 223 complete responses. To ensure alignment with students' coursework and tasks, the study was limited to Mechanical and Aerospace Engineering students. The questionnaire comprised 17 items across four constructs. For item-feature construction, we gathered data on assigned WIL opportunities, including job settings and tasks (Table 1). To model user features, we collected data on their coursework history, major, background, and engineering self-efficacy lev-
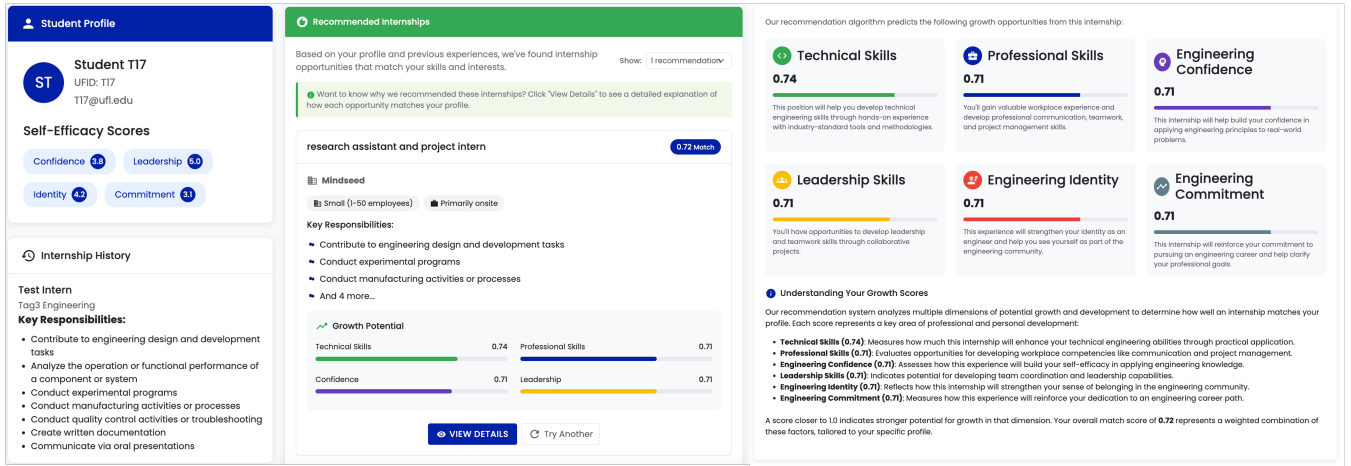
Figure 2: Interface of our WIL opportunity Recommender System, Pro-CaRE

Table 1: Summary of Participant Characteristics

Category	N	%	Category	N	%
School Year			**Gender**		
1st year	47	21.1%	Female	80	35.8%
2nd year	69	30.9%	Male	143	64.1%
3rd year	80	35.9%	**Major**		
4th year	24	10.8%	Mechanical	119	53.3%
5th year	3	1.4%	Aerospace	99	44.3%
			Missing	5	2.2%
Company Size			**Setting**		
Small	31	13.9%	Onsite	191	85.6%
Medium	20	9.0%	Hybrid	22	9.8%
Large	172	77.1%	Remote	10	4.4%

els. Lastly, rating or interaction data was obtained by collecting students' self-perceived learning outcomes.

3.5 Stage 1: Identifying Task Components

Table 2 presents the descriptive statistics of tasks in WIL experiences reported by the participants, which serve as the recommender system's item features. On average, students reported engaging in 6.55 tasks during their WIL experiences, with the number of tasks ranging from a minimum of 1 to a maximum of 15. A similar process was applied to a dataset of job postings we collected. We extracted a total of 1,913 specific job tasks from the internship postings, categorizing them into 15 predefined groups. First, we extracted the description sections of aerospace internship postings. Then, we used Llama 2 to summarize these descriptions into a list of specific responsibilities and tasks. This process yielded 1,913 unique task entities, which were subsequently labeled by human annotators. Pre-trained models BERT, RoBERTa, and DeBERTa were fine-tuned using the AutoModel class from the Hugging Face Transformers library to classify the engineering tasks into the 15 predefined categories. The models were trained using a system equipped with an NVIDIA GeForce RTX 3060 GPU. The output of this stage was the item feature vectors associated with the jobs provided by our participants (along with the learning outcomes) and the newly processed vectors from the

unseen recommendations (LLM-processed job postings).

3.6 Stage 2: System Implementation

Our LLM-enhanced hybrid content-based recommender system provides personalized internship recommendations based on six outcome variables: two learning outcomes (technical and professional skills) and four engineering self-efficacy measures (confidence, leadership, identity, and commitment). To compute recommendations, the system first constructs distance matrices using Euclidean distance for both users and internships. These matrices measure similarity, with diagonal values set to infinity to prevent self-matching. The closest N users and internships are identified, allowing the system to suggest opportunities based on student profiles and past participant preferences. For students with no prior internship interactions, missing outcome scores are imputed using the average of corresponding variables. This enriched dataset is then used to train the recommendation model. Truncated SVD is applied to a sparse interaction matrix, where users (rows), internships (columns), and outcome scores (values) represent engagement data. Additionally, for jobs that have not been introduced to the user—unseen job postings—the system utilizes LLM-processed descriptions to generate feature representations. These representations enable the model to infer and recommend relevant internship opportunities even when no direct interaction data exists.

3.7 Stage 3: Evaluating the Efficacy

In order to internally evaluate the system's efficacy and address our **RQ2**, our original dataset was split 70% for training and 30% for testing. After training, the reconstructed test matrix predicts scores for unobserved user-internship pairs. Model performance is evaluated using mean squared error (MSE) and R^2 scores, where lower MSE and higher R^2 indicate better accuracy. The system evaluated different values for top n (the number of closest users or items considered) from 1 to 5. The R^2 scores across all outcomes consistently show values above 0.72, suggesting that the model explains a substantial proportion of the variance in internship recommendation scores. The MSE values, which measure prediction error, remain low, reinforcing the model's accuracy. Notably, as top n increases, the R^2 values tend to rise, peaking

Table 2: WIL Opportunities and Relevant Tasks - Item Features

Task Description	Student Response		LLM-processed WIL Postings
	Freq	Work Hours (%)	Freq
Participate in on-the-job training and develop new technical skills	181	20.42%	127
Collaborate with an engineering group, department, multi-disciplinary team	166	28.40%	196
Contribute to engineering design and development tasks	150	29.77%	219
Create written documentation of procedures, processes, or results	149	21.14%	171
Analyze the operation or functional performance of a component/system	143	24.11%	74
Communicate via oral presentations to a variety of audiences	143	10.65%	86
Conduct quality control activities or troubleshoot a failure of a component or system	96	17.82%	83
Conduct manufacturing activities or processes	88	22.22%	83
Develop or modify computer codes and/or public software or computational tools	87	27.72%	169
Conduct experimental programs, prototypes, components, hardware, or products	86	19.88%	166
Create or revise technical drawings	84	18.19%	37
Perform thermal science or fluid dynamic analysis	35	17.70%	35
Perform solid mechanics analysis	32	14.38%	31
Perform dynamics or vibrational analysis	21	14.84%	22
Perform control analysis	9	9.67%	47

around top n = 4, indicating that incorporating a larger but limited number of similar users/items enhances predictive power. For example, for technical skills, the MSE decreases from 0.106 (top n=5) to 0.063 (top n=1), while the R^2 value increases from 0.793 (top n=5) to 0.734 (top n=1). A similar pattern is observed across the other outcome variables, suggesting that a moderate number of closest users/items yields optimal recommendations.

3.8 Stage 4: Evaluating the Explainability

As shown in Figure 2, we have incorporated both local and global approaches to explainability measures [7], addressing our **RQ3**, and focusing on the explainability of our recommendations. This ongoing phase establishes how common approaches to implementing explainable AI in educational settings—implemented with General Explanation (explaining why a specific item is recommended), Feature Relevance/Importance, and Example-based methods (selecting particular instances to explain the model)—can help students build trust in and perceive relevance in the recommendation outputs. We are currently working on collecting qualitative data through interviews with students to evaluate the efficacy of these added features in our system.

4. CONCLUSION

In this study, we introduced an LLM-enhanced hybrid content-based recommender system to provide systematic support for WIL opportunities for engineering students. We have demonstrated, in response to RQ1, how a scalable recommender system is designed and implemented to produce recommendation outputs for engineering students. Our internal evaluation indicates promising performance in predicting students' growth in core learning outcomes based on our recommendations, addressing RQ2. The ongoing investigation into its explainability features (RQ3) will shed more light on the efficacy and relevance of the system in greater depth.

5. ACKNOWLEDGMENT

6. REFERENCES

[1] P. Ackerman and K. Arcieri. Co-ops are great! but what are the numbers telling us? In *2022 ASEE Annual Conference & Exposition*, 2022.

[2] S. R. Brunhaver, R. F. Korte, S. R. Barley, and S. D. Sheppard. Bridging the gaps between engineering education and practice. In *US engineering in a global economy*, pages 129–163. University of Chicago Press, 2017.

[3] H. Drachsler. *Trusted learning analytics*. Universität Hamburg, 2018.

[4] K. Fabian, E. Taylor-Smith, S. Smith, D. Meharg, and A. Varey. An exploration of degree apprentice perspectives: a q methodology study. *Studies in Higher Education*, 47(7):1397–1409, 2022.

[5] S. G. Gebeyehu and E. B. Atanaw. Impact of internship program on engineering and technology education in ethiopia: Employers' perspective. *Journal of Education and Training*, 5(2):64–68, 2018.

[6] E. C. Johnson, M. Villafañe-Delgado, D. Symonette, K.-A. Carr, M. Hughes, J. Burroughs, S. Floryanzia, M. Cervantes, and W. R. Gray-Roncal. An immersive curriculum to develop computational science and research skills in a cohort-based internship program. In *2023 IEEE Integrated STEM Education Conference (ISEC)*, pages 293–300. IEEE, 2023.

[7] H. Khosravi, S. B. Shum, G. Chen, C. Conati, Y.-S. Tsai, J. Kay, S. Knight, R. Martinez-Maldonado, S. Sadiq, and D. Gašević. Explainable artificial intelligence in education. *Computers and education: artificial intelligence*, 3:100074, 2022.

[8] D. Kurniadi, E. Abdurachman, H. Warnars, and W. Suparta. A proposed framework in an intelligent recommender system for the college student. In *Journal of Physics: Conference Series*, volume 1402, page 066100. IOP Publishing, 2019.

[9] J. B. Main, A. L. Griffith, X. Xu, and A. M. Dukes. Choosing an engineering major: A conceptual model of student pathways into engineering. *Journal of Engineering Education*, 111(1):40–64, 2022.

[10] R.-I. Mogoş and C.-N. Bodea. Recommender systems for engineering education. *Rev. Roum. Sci. Techn.–Électrotechn. et Énerg*, 64(4):435–442, 2019.

[11] E. Mousavinasab, N. Zarifsanaiey, S. R. Niakan Kalhori, M. Rakhshan, L. Keikha, and M. Ghazi Saeedi. Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, 29(1):142–163, 2021.

[12] J. Muangprathub, V. Boonjing, and K. Chamnongthai. Learning recommendation with formal concept analysis for intelligent tutoring system. *Heliyon*, 6(10), 2020.

[13] P. Sattler, R. Wiggers, and C. Arnold. Combining workplace training with postsecondary education: The spectrum of work-integrated learning (wil) opportunities from apprenticeship to experiential learning. *Canadian Apprenticeship Journal*, 5:1–33, 2011.

[14] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen. Collaborative filtering recommender systems. In *The adaptive web: methods and strategies of web personalization*, pages 291–324. Springer, 2007.

[15] M. C. Urdaneta-Ponte, A. Mendez-Zorrilla, and I. Oleagordia-Ruiz. Recommendation systems for education: Systematic review. *Electronics*, 10(14):1611, 2021.

[16] S. M. Zehr and R. Korte. Student internship experiences: learning about the workplace. *Education+ Training*, 62(3):311–324, 2020.