VẬT LIỆU & XÂY DỰNG
Tạp chí Vật liệu và Xây dựng
ISSN 1859 - 381X
Website: www. jomc.vn

Đánh giá khả năng ứng dụng của AI trong quản lý hợp đồng xây dựng

Nguyễn Trọng Phúc[1], Phạm Trương Thị Lệ Hiểu[1*], Hà Duy Khánh[1]

[1] Khoa Xây dựng, Trường Đại học Công nghệ Kỹ thuật Thành phố Hồ Chí Minh

TỪ KHOÁ

Quản lý hợp đồng xây dựng
Trí tuệ nhân tạo
Xử lý ngôn ngữ tự nhiên
Xử lý văn bản
BERT
Mô hình ngôn ngữ lớn

TÓM TẮT

Xử lý ngôn ngữ tự nhiên (Natural Language Processing - NLP) ngày càng được ứng dụng rộng rãi trong quản lý hợp đồng xây dựng. Tuy nhiên, các nghiên cứu hiện nay chưa khảo sát và so sánh một cách hệ thống ảnh hưởng của các yếu tố như sử dụng mô hình đã được tiền huấn luyện (pre-trained model), kiến trúc đơn nhiệm hay đa nhiệm, cũng như mối quan hệ giữa các yếu tố này, quy mô dữ liệu và hiệu suất mô hình (accuracy, F1 score, tỷ lệ hài lòng và tỷ lệ thành công). Khoảng trống này làm hạn chế cơ sở tham khảo định lượng cho việc lựa chọn kiến trúc mô hình và chiến lược huấn luyện phù hợp với từng bài toán NLP trong quản lý hợp đồng xây dựng. Để giải quyết những giới hạn này, nghiên cứu này tổng quan có hệ thống 124 bài báo về ứng dụng NLP trong quản lý hợp đồng xây dựng, phân loại theo nhóm tác vụ, loại tài liệu và phương pháp, đồng thời xem xét ảnh hưởng của mô hình tiền huấn luyện, quy mô dữ liệu huấn luyện và kiến trúc đơn nhiệm/đa nhiệm đến hiệu suất mô hình. Kết quả cho thấy phân loại văn bản là nhóm tác vụ phổ biến nhất (60%), điều khoản hợp đồng là loại tài liệu được nghiên cứu nhiều nhất (46%), trong khi các mô hình học máy (Machine Learning - ML) chiếm tỷ lệ cao nhất (26%) và Bidirectional Encoder Representations from Transformers (BERT) cùng các biến thể đạt hiệu suất trung bình cao nhất (89%). Các mô hình tiền huấn luyện có lợi thế rõ rệt khi dữ liệu gán nhãn hạn chế, trong khi mô hình đa nhiệm cho thấy hiệu quả tốt hơn ở quy mô dữ liệu trung bình. Bên cạnh đó, nghiên cứu đồng thời chỉ ra các hướng nghiên cứu nhằm phát triển mô hình ngôn ngữ chuyên ngành, bộ dữ liệu chuẩn hóa và công khai, dữ liệu hợp đồng xây dựng tiếng Việt, cũng như các kiến trúc đa nhiệm để nâng cao khả năng ứng dụng thực tiễn.

KEYWORDS

Construction contract management
Artificial intelligence
Natural language processing
Text processing
BERT
Large language models

ABSTRACT

Natural language processing (NLP) has been increasingly applied to construction contract management. However, existing studies have not systematically investigated the effects of key factors, including the use of pre-trained models, single-task versus multi-task models, and the interactions among these factors, training data size, and model performance (e.g., accuracy, F1 score, satisfaction rate, and success rate). This gap limits the quantitative evidence to guide the selection of appropriate model architectures and training strategies for specific NLP tasks in construction contract management. To address these limitations, this study systematically reviewed 124 papers on NLP applications in construction contract management and classified them according to task types, document types, and methodological approaches. This study further examined the effects of pre-trained models, training dataset size, and single-task/multi-task architectures on model performance. The results show that text classification was the most prevalent task type (60%), contract clauses were the most extensively studied document type (46%), and traditional machine learning models accounted for the largest proportion of the applied approaches (26%). Among the applied approaches, Bidirectional Encoder Representations from Transformers (BERT) and its variants achieved the highest average performance (89%). Pre-trained models demonstrated clear advantages when labeled training data are limited, whereas multi-task architectures tended to perform more effectively with moderate-sized datasets. Furthermore, this study identified several research directions, including the development of domain-specific language models, standardized and publicly available benchmark datasets, Vietnamese construction contract datasets, and multi-task architectures to enhance the practical application of NLP in construction contract management.

1. Giới thiệu và tổng quan nghiên cứu

Hợp đồng xây dựng là công cụ pháp lý quan trọng, quy định quyền và nghĩa vụ của chủ đầu tư, nhà thầu và các bên liên quan trong suốt vòng đời dự án, từ giai đoạn đấu thầu đến thi công và nghiệm thu. Tuy nhiên, do đặc thù các dự án xây dựng thường có thời gian thực

*Liên hệ tác giả: hieupttl@hcmute.edu.vn
Nhận ngày 04/09/2026, sửa xong ngày 24/09/2026, chấp nhận đăng ngày 25/09/2026
Link DOI: https://doi.org/10.54772/jomc.05.2026.1508

hiện dài, quy mô lớn, tính phức tạp cao và sự tham gia của nhiều bên, hợp đồng xây dựng tiềm ẩn nhiều rủi ro như điều khoản không rõ ràng, thiếu sót, mâu thuẫn giữa các tài liệu hoặc thay đổi phạm vi công việc trong quá trình thực hiện. Những vấn đề này có thể dẫn đến khiếu nại, tranh chấp và kiện tụng, gây tổn thất đáng kể về chi phí và thời gian cho các bên tham gia cũng như toàn ngành xây dựng [1]. Mức độ nghiêm trọng của vấn đề được thể hiện qua các báo cáo của HKA. Theo Fifth Annual CRUX Insight Report công bố năm 2022, tổng giá trị các khoản tiền đang tranh chấp vượt 80 tỷ USD, trong khi tổng thời gian gia hạn tiến độ được yêu cầu lên tới hơn 840 năm [2]. Một phân tích khác của HKA, dựa trên hơn 2.000 tỷ USD giá trị các dự án được thực hiện trong giai đoạn từ tháng 3/2020 đến tháng 8/2021, cũng ghi nhận tổng giá trị tranh chấp lên tới 73 tỷ USD và tổng thời gian gia hạn tiến độ được yêu cầu tương đương 750 năm dự án [3]. Những con số này cho thấy các vấn đề phát sinh từ hợp đồng có thể tác động đáng kể đến chi phí và tiến độ thực hiện dự án, qua đó khẳng định quản lý hợp đồng xây dựng không chỉ là hoạt động mang tính hành chính - pháp lý mà còn giữ vai trò quan trọng trong kiểm soát chi phí, tiến độ và rủi ro của dự án, đòi hỏi quá trình quản lý và rà soát hợp đồng phải được thực hiện một cách chính xác, kịp thời và có hệ thống.

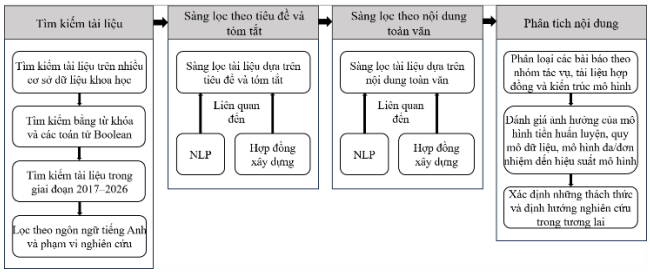
Trong thực tiễn cũng như trong phần lớn các nghiên cứu trước đây, việc quản lý hợp đồng xây dựng - bao gồm rà soát điều khoản, nhận diện rủi ro và phát hiện mâu thuẫn - chủ yếu vẫn được thực hiện theo phương pháp thủ công: các chuyên gia pháp lý, kỹ sư hợp đồng đọc toàn bộ văn bản, đối chiếu giữa các tài liệu (điều kiện hợp đồng, phụ lục, hồ sơ mời thầu, bản vẽ kỹ thuật) và chỉnh sửa dựa trên kinh nghiệm cá nhân. Cách tiếp cận này gặp nhiều hạn chế đáng kể do khối lượng văn bản hợp đồng xây dựng thường rất lớn và ngôn ngữ pháp lý - kỹ thuật phức tạp, khiến việc rà soát thủ công tốn nhiều thời gian, nhân lực, dễ bỏ sót các điều khoản rủi ro, đồng thời đòi hỏi kinh nghiệm và chuyên môn cao từ người thực hiện [4]. Trong bối cảnh áp lực về tiến độ đấu thầu và thi công ngày càng lớn, việc rà soát toàn bộ văn bản hợp đồng trong thời gian hạn chế theo cách thủ công gần như là bất khả thi, đặc biệt khi khối lượng thông tin cần đối chiếu liên quan đến rủi ro và trách nhiệm các bên là rất lớn [5]. Điều này đặt ra yêu cầu cấp thiết phải phát triển một phương pháp quản lý hợp đồng xây dựng tự động, nhanh chóng và chính xác hơn, có khả năng xử lý khối lượng văn bản lớn mà vẫn đảm bảo độ tin cậy tương đương hoặc vượt trội so với con người. Trong bối cảnh đó, trí tuệ nhân tạo (artificial intelligence - AI) và xử lý ngôn ngữ tự nhiên (Natural language processing - NLP) đã ra đời và phát triển mạnh mẽ nhằm đáp ứng nhu cầu này, nhờ khả năng tự động đọc hiểu, phân loại và trích xuất thông tin từ dữ liệu văn bản của hợp đồng với tốc độ và quy mô vượt xa khả năng xử lý thủ công của con người [6].

Nhận thấy tiềm năng đó, gần đây đã xuất hiện một số nghiên cứu tổng quan về ứng dụng NLP trong quản lý hợp đồng xây dựng. Cụ thể, Soy và cộng sự [7] đã thực hiện một tổng quan hệ thống dựa trên cơ sở dữ liệu Scopus, sàng lọc các bài báo liên quan đến việc ứng dụng NLP hỗ trợ phân tích hợp đồng xây dựng, tập trung mô tả phương pháp và kết quả đạt được của từng nghiên cứu. Một nghiên cứu khác được thực hiện bởi Uddin và cộng sự [8] đã khảo sát diện rộng hơn các mô hình NLP được áp dụng cho nhiều nhóm điều khoản khác nhau, từ điều khoản hợp đồng chung đến hồ sơ tranh chấp. Gần đây nhất, một nghiên cứu bởi Aktas và cộng sự [9] đã hệ thống hóa các xu hướng nghiên cứu, phạm vi ứng dụng (phân tích hợp đồng, quản lý rủi ro, số hóa và tích hợp quản lý dự án) và những thách thức khi triển khai AI trong lĩnh vực hợp đồng xây dựng. Tuy nhiên, điểm chung của các nghiên cứu tổng quan nêu trên là chủ yếu dừng lại ở mức thống kê mô tả - liệt kê mô hình nào được sử dụng cho bài toán nào và đạt hiệu suất (accuracy, F1 score, tỷ lệ hài lòng và tỷ lệ thành công) bao nhiêu - mà chưa đi sâu khảo sát, so sánh hiệu suất của các mô hình dưới góc độ những yếu tố kỹ thuật quan trọng như: việc sử dụng mô hình tiền huấn luyện hay huấn luyện từ đầu, kiến trúc đơn nhiệm (single-task) hay đa nhiệm (multi-task), và đặc biệt là mối quan hệ giữa các yếu tố này và quy mô dữ liệu với hiệu suất mô hình. Những khoảng trống này khiến các nhà nghiên cứu và người thực hành thiếu một cơ sở tham khảo định lượng, có hệ thống để lựa chọn kiến trúc mô hình và chiến lược đào tạo phù hợp khi triển khai một bài toán NLP cụ thể trong quản lý hợp đồng xây dựng. Để giải quyết vấn đề này, nghiên cứu này đề xuất khảo sát có hệ thống 124 bài báo khoa học về ứng dụng NLP trong quản lý hợp đồng xây dựng thông qua việc phân loại theo nhóm tác vụ, loại tài liệu, phương pháp và còn xét đến ảnh hưởng của việc sử dụng mô hình tiền huấn luyện, quy mô dữ liệu huấn luyện, và kiến trúc đơn nhiệm/đa nhiệm đến hiệu suất mô hình, nhằm cung cấp cơ sở tham khảo thực tiễn cho việc lựa chọn kiến trúc AI và chiến lược đào tạo phù hợp trong các nghiên cứu và ứng dụng NLP trong quản lý hợp đồng xây dựng.

Cấu trúc của bài báo được trình bày như sau: Mục 2 trình bày phương pháp nghiên cứu, bao gồm quy trình thu thập, sàng lọc và phân tích nội dung của 124 bài báo được lựa chọn. Mục 3 trình bày kết quả phân loại các nghiên cứu theo ba nhóm tác vụ NLP, gồm phân loại văn bản, trích xuất thông tin và tạo sinh văn bản; đồng thời phân tích loại tài liệu và điều khoản hợp đồng được nghiên cứu, các phương pháp NLP được áp dụng và phương pháp biểu diễn văn bản. Mục 4 phân tích các yếu tố ảnh hưởng đến hiệu suất của các mô hình NLP, tập trung vào việc sử dụng mô hình tiền huấn luyện, quy mô dữ liệu huấn luyện và kiến trúc mô hình đơn nhiệm hoặc đa nhiệm. Mục 5 thảo luận những thách thức còn tồn tại và đề xuất các hướng nghiên cứu tiềm năng trong tương lai. Cuối cùng, Mục 6 tổng kết những phát hiện chính và các đóng góp của nghiên cứu.

2. Phương pháp nghiên cứu

Nghiên cứu này áp dụng phương pháp tổng quan dựa trên phân tích nội dung (content analysis-based review) - một phương pháp được sử dụng phổ biến để tổng hợp, hệ thống hóa và trích xuất tri thức từ tài liệu trong lĩnh vực kỹ thuật và quản lý xây dựng. Quy trình tổng quan được thực hiện qua bốn giai đoạn liên tiếp (Hình 1): (1) tìm kiếm tài liệu, (2) sàng lọc theo tiêu đề và tóm tắt, (3) sàng lọc theo nội dung toàn văn, và (4) phân tích nội dung theo góc độ tác vụ, kiến trúc mô hình, phạm vi ứng dụng và hiệu suất.



Hình 1. Phương pháp nghiên cứu.

2.1. Tìm kiếm tài liệu

Quá trình tìm kiếm tài liệu được thực hiện trên nhiều cơ sở dữ liệu khoa học, bao gồm ScienceDirect, ASCE Library, IEEE Xplore, MDPI, arXiv, SpringerLink và Google Scholar. Các từ khóa kết hợp với toán tử Boolean AND/OR được sử dụng để đảm bảo bao quát tài liệu liên quan, với chuỗi tìm kiếm dạng: ("artificial intelligence" OR "machine learning" OR "deep learning" OR "natural language processing" OR "large language model") AND ("construction contract" OR "construction claim" OR "construction dispute" OR "contract management" OR "construction specification"). Phạm vi tìm kiếm giới hạn ở các bài báo khoa học công bố bằng tiếng Anh trong giai đoạn từ năm 2017 đến năm 2026. Các bài báo trùng lặp được xác định và loại bỏ dựa trên tiêu đề. Sau quá trình sàng lọc ban đầu, tổng cộng 386 bài báo được giữ lại để tiếp tục thực hiện các bước sàng lọc tiếp theo.

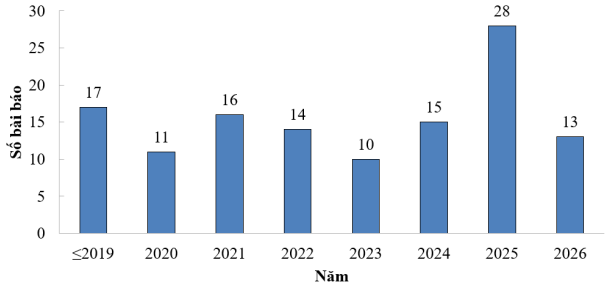
2.2. Sàng lọc theo tiêu đề và tóm tắt

Giai đoạn này nhằm xác định các nghiên cứu có liên quan trực tiếp đến lĩnh vực quản lý hợp đồng xây dựng và ứng dụng NLP. Các tài liệu thu được từ bước tìm kiếm được rà soát thủ công thông qua tiêu đề và tóm tắt; những bài báo không đề cập đến các khái niệm liên quan đến hợp đồng xây dựng hoặc không ứng dụng NLP đều bị loại bỏ khỏi danh sách. Sau quá trình sàng lọc này, tổng cộng 168 bài báo được giữ lại để tiếp tục thực hiện các bước sàng lọc tiếp theo.

2.3. Sàng lọc theo nội dung toàn văn

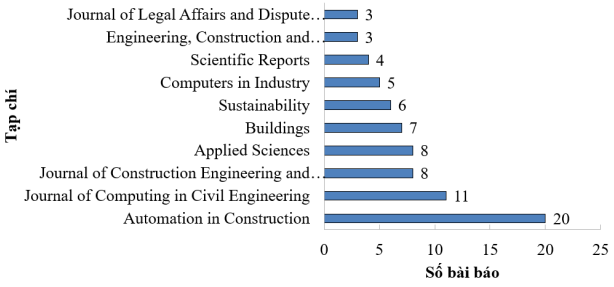
Các tài liệu còn lại sau bước sàng lọc trước tiếp tục được kiểm tra bằng cách đọc toàn văn, nhằm loại bỏ những nghiên cứu tuy có chứa các từ khóa liên quan trong tiêu đề hoặc tóm tắt nhưng không thực sự tập trung vào ứng dụng NLP trong quản lý hợp đồng xây dựng (ví dụ: các bài chỉ đề cập đến "hợp đồng" hoặc "xây dựng" nhưng không ứng dụng AI/NLP). Sau ba vòng sàng lọc liên tiếp, 124 bài báo được giữ lại để tiến hành phân tích chuyên sâu.

Hình 2 cho thấy bức tranh khá rõ nét về sự phân bố 124 bài báo nhằm thể hiện sự phát triển của hướng nghiên cứu NLP qua thời gian. Cụ thể, số lượng công bố về ứng dụng NLP trong quản lý hợp đồng xây dựng tăng trưởng rõ rệt trong những năm gần đây. Giai đoạn từ năm 2019 trở về trước chỉ ghi nhận 17 bài báo, trong khi riêng năm 2025 đã có 28 bài báo được công bố và cũng là năm có số lượng công bố cao nhất trong toàn bộ giai đoạn khảo sát. Tính chung 4 năm gần nhất (2023–2026), số bài báo đạt 66 bài, phản ánh sự tăng tốc nghiên cứu song hành với sự phát triển của các mô hình ngôn ngữ lớn (large language model - LLM) như generative pre-trained transformer 4 (GPT-4), GPT-4o và large language model meta AI 3 (LLaMA-3). Nhìn chung, xu hướng tăng dần qua các năm khẳng định đây là một hướng nghiên cứu mới nổi, đang thu hút sự quan tâm ngày càng mạnh mẽ của cộng đồng khoa học trong lĩnh vực kỹ thuật xây dựng.



Hình 2. Sự phân bố các nghiên cứu ứng dụng NLP trong quản lý hợp đồng xây dựng theo thời gian.

Xét về loại tạp chí, Hình 3 cho thấy Automation in Construction là tạp chí dẫn đầu với 20 bài báo, tiếp theo là Journal of Computing in Civil Engineering với 11 bài báo. Journal of Construction Engineering and Management và Applied Sciences đồng hạng ba với 8 bài báo. Đáng chú ý, phần lớn các tạp chí có số lượng công bố cao đều thuộc các nhà xuất bản uy tín trong lĩnh vực kỹ thuật xây dựng như Elsevier và American Society of Civil Engineers (ASCE), cho thấy độ tin cậy học thuật cao của bộ dữ liệu được thu thập.



Hình 3. Top 10 tạp chí có số lượng phân bố bài báo nhiều nhất.

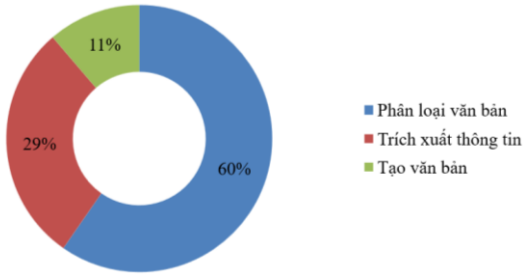
2.4. Phân tích nội dung

Phân tích nội dung được thực hiện dựa trên năm khía cạnh chính: (1) phân loại các bài báo theo nhóm tác vụ NLP, gồm phân loại văn bản, trích xuất thông tin và tạo sinh văn bản; (2) phân loại các nhóm tài liệu/điều khoản hợp đồng, gồm điều khoản hợp đồng, điều luật/văn bản pháp lý, hồ sơ mời thầu/đấu thầu, hồ sơ chi phí/dự toán và hồ sơ tranh chấp/khiếu nại; (3) tổng hợp kiến trúc mô hình được sử dụng,

gồm học máy truyền thống, BERT và các biến thể, học sâu, rule-based/NLP truyền thống và LLM/GPT/LLaMA/T5; (4) phân tích các phương pháp biểu diễn văn bản (text representation) được áp dụng; và (5) đánh giá ảnh hưởng của việc sử dụng mô hình tiền huấn luyện, quy mô dữ liệu huấn luyện, cũng như loại mô hình (mô hình đa nhiệm và đơn nhiệm) đến hiệu suất của các mô hình. Do dữ liệu gốc được thu thập dưới dạng văn bản tự do, nhóm nghiên cứu đã chuẩn hóa và phân nhóm bằng từ khóa và nội dung để tổng hợp thống kê; chỉ những bài báo có công bố hiệu suất định lượng rõ ràng mới được đưa vào tính giá trị trung bình và so sánh giữa các nhóm.

3. Ứng dụng NLP trong quản lý hợp đồng xây dựng

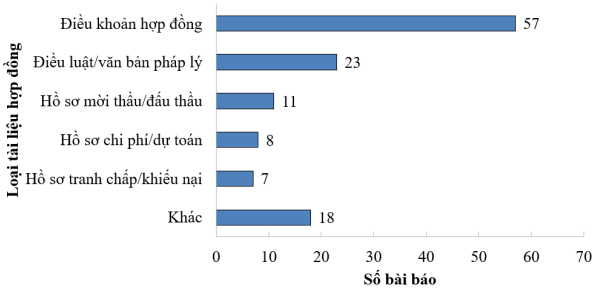
Toàn bộ 124 bài báo được phân loại theo 3 nhóm tác vụ NLP chính như được thể hiện ở Hình 4: (1) phân loại văn bản - nhận diện, gán nhãn rủi ro, phân loại điều khoản; (2) trích xuất thông tin - trích xuất thực thể, quan hệ, điều khoản, dữ liệu có cấu trúc từ văn bản; (3) tạo sinh văn bản - soạn thảo, tóm tắt, trả lời câu hỏi tự động. Kết quả cho thấy phân loại văn bản là tác vụ chủ đạo với 60 %, phù hợp với nhu cầu thực tiễn phổ biến nhất trong quản lý hợp đồng xây dựng là nhận diện điều khoản rủi ro, phân loại điều khoản, dự đoán hành động xử lý rủi ro. Trích xuất thông tin đứng thứ hai với 29 %, phục vụ cho trích xuất nghĩa vụ, phát hiện xung đột điều khoản và xây dựng cơ sở dữ liệu hợp đồng có cấu trúc. Tạo sinh văn bản chỉ chiếm 11 %, gắn liền với sự xuất hiện gần đây của các LLM (GPT-4o, LLaMA-3) có khả năng sinh văn bản, hướng đến mục tiêu tự động soạn thảo và phản hồi các nội dung liên quan đến hồ sơ thầu, hợp đồng.



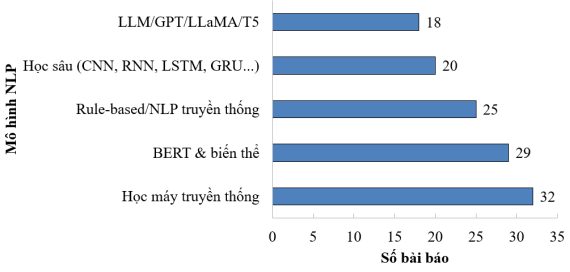
Hình 4. Tỷ lệ phân bố các bài báo vào các nhóm tác vụ chính.

Về loại tài liệu/điều khoản được các nghiên cứu khai thác, Hình 5 cho thấy nhóm "điều khoản hợp đồng" chiếm tỷ trọng cao nhất với 57 bài báo. Theo sau là nhóm "điều luật/văn bản pháp lý" với 23 bài báo. Nhóm "hồ sơ mời thầu/đấu thầu" gồm 11 bài báo và "hồ sơ chi phí/dự toán" gồm 8 bài báo. Đáng chú ý, nhóm "hồ sơ tranh chấp/khiếu nại" chiếm tỷ trọng thấp nhất với chỉ 7 bài báo. Tỷ trọng thấp của nhóm "hồ sơ tranh chấp/khiếu nại" này so với các loại tài liệu/điều khoản khác cho thấy đây vẫn là một khoảng trống nghiên cứu đáng được khai thác trong các nghiên cứu tiếp theo.

Xét theo kiến trúc mô hình được sử dụng, Hình 6 cho thấy nhóm học máy truyền thống (Support Vector Machine - SVM, Random Forest, XGBoost, k-Nearest Neighbors - kNN và Decision Tree) chiếm ưu thế với 32 bài báo, nhờ tính đơn giản, dễ triển khai và phù hợp với các bộ dữ liệu quy mô nhỏ vốn phổ biến trong lĩnh vực hợp đồng xây dựng. Bidirectional Encoder Representations from Transformers (BERT) và các biến thể (RoBERTa, DistilBERT, LegalBERT, SciBERT và DeBERTa) đứng thứ hai với 29 bài báo, nhờ khả năng sử dụng các mô hình tiền huấn luyện linh hoạt cho dữ liệu chuyên ngành xây dựng với chi phí tính toán vừa phải. Nhóm rule-based/NLP truyền thống chiếm 25 bài báo, phản ánh việc nhiều nghiên cứu giai đoạn đầu vẫn dựa vào cách tiếp cận thủ công/bán tự động. Nhóm học sâu (Convolutional Neural Network - CNN, Recurrent Neural Network - RNN, Long Short-Term Memory – LSTM và Gated Recurrent Unit - GRU) chiếm 20 bài báo. Nhóm LLM/GPT/LLaMA/T5 xếp cuối với 18 bài báo nhưng gần như toàn bộ tập trung trong giai đoạn 2023–2026, cho thấy đây là làn sóng công nghệ mới nhất đang được cộng đồng nghiên cứu tích cực khai thác và dự kiến sẽ tiếp tục tăng trưởng trong các năm tới.

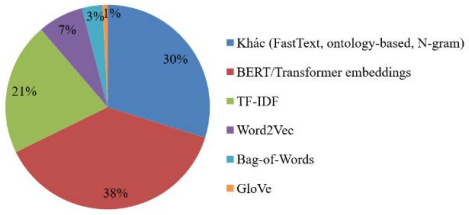


Hình 5. Phân bố loại tài liệu/điều khoản hợp đồng được nghiên cứu.



Hình 6. Mô hình NLP phổ biến trong quản lý hợp đồng xây dựng.

Về phương pháp biểu diễn văn bản, Hình 7 cho thấy BERT/Transformer embeddings chiếm tỷ trọng lớn nhất với 38 % nhờ khả năng nắm bắt ngữ nghĩa sâu, phù hợp với văn bản pháp lý hợp đồng vốn có cấu trúc câu phức tạp. Nhóm các phương pháp khác (gồm FastText, ontology-based và N-gram) đứng thứ hai với 30 % do tính đa dạng cao - mỗi nghiên cứu thường tùy biến phương pháp biểu diễn theo đặc thù bài toán cụ thể. TF-IDF vẫn được ưa chuộng cho các bài toán phân loại truyền thống, chiếm 21 %, nhờ tính đơn giản, dễ áp dụng. Word2Vec chiếm 7 % nhờ cân bằng tốt giữa hiệu quả ngữ nghĩa và chi phí tính toán. Còn lại các phương pháp khác như Bag-of-Words và GloVe chiếm tỷ lệ thấp hơn với 3 % và 1 % do đạt hiệu suất thấp hơn trong việc trích xuất các ngữ nghĩa đặc trưng của hợp đồng xây dựng.



Hình 7. Phân bố tỷ lệ các phương pháp biểu diễn văn bản.

3.1. Ứng dụng NLP trong phân loại văn bản của hợp đồng xây dựng

Nhìn chung, các nghiên cứu cho thấy NLP đang chuyển phân tích hợp đồng xây dựng từ rà soát thủ công sang tự động phân loại và sàng lọc nội dung, với dữ liệu đầu vào chủ yếu là hợp đồng, điều khoản hợp đồng, hồ sơ pháp lý và các tài liệu liên quan. Các bài toán phổ biến gồm phân loại điều khoản rủi ro, xác định mức độ rủi ro, phân loại trách nhiệm/các bên liên quan, kiểm tra tuân thủ, tìm điều khoản tương đồng, phân loại nội dung chuyên môn và phát hiện điều khoản thiếu (Bảng 1). Các mô hình từ SVM, Random Forest, CNN, RNN đến BERT và LLM thường đạt hiệu suất khoảng 70-100 %. Những kết quả này cho thấy NLP có tiềm năng giảm thời gian đọc và rà soát hàng nghìn điều khoản, hỗ trợ chuyên gia nhanh chóng xác định các nội dung bất thường hoặc có rủi ro, đồng thời tạo nền tảng cho các hệ thống tự động rà soát, đánh giá rủi ro và hỗ trợ ra quyết định trong quản lý hợp đồng xây dựng.

Bảng 1. Các nghiên cứu về ứng dụng NLP trong phân loại văn bản của hợp đồng xây dựng.

Mô hình	Tiền huấn luyện	Mục tiêu	Hiệu suất	Số lượng data	Trích dẫn
MatchPyramid, BERT	Có	Phát hiện điều khoản bị thiếu	93 % (Accuracy)	1.322	[10]
BERT	Có	Nhận diện điều khoản rủi ro, phân loại mức độ rủi ro	93 % (F1 score)	2.807	[11]
BERT	Có	Xác định loại rủi ro, bên chịu trách nhiệm và hình thức xử lý rủi ro	90 % (F1 score)	2.586	[6]
Temporal convolutional network (TCN)	Không	Phân loại các tiêu chuẩn đo lường chi phí xây dựng	90 % (F1 score)	50.000	[12]
BERT	Có	Danh mục rủi ro	86 % (F1 score)	1.000	[13]
RNN	Có	Các hạng mục	93 % (F1 score)	3.000	[14]
SVM	Không	Danh mục kiểm tra tuân thủ	100 % (Accuracy)	330	[15]
XGBOOST	Không	Phân loại điều khoản yêu cầu báo cáo	99 % (F1 score)	8.943	[16]
BERT	Có	Phân loại rủi ro, loại câu, xác định bên liên quan	73 % (F1 score)	3.500	[17]
GPT-2	Có	Phân loại điều khoản rủi ro	81 %	3.411	[18]
BERT	Có	Phân loại rủi ro hợp đồng, phát hiện điều khoản tương đồng, phát hiện bất thường	92 % (Accuracy)	20.000	[19]
BERT	Có	Phân loại câu hợp đồng có rủi ro	89 % (Accuracy)	3.485	[20]
MLP	Không	Phân loại tranh chấp	87 % (F1 score)	375	[21]
SVM	Không	Phân loại câu hợp đồng vào các nhóm thời gian, chi phí và chất lượng	88 % (F1 score)	827	[22]
BERT	Có	Phân loại điều khoản	77 % (F1 score)	733	[23]
BERT	Có	Phân loại điều khoản	98 % (F1 score)	1.277	[24]

Ứng dụng NLP trong phân loại văn bản hợp đồng xây dựng chủ yếu tập trung vào phân loại và đánh giá rủi ro hợp đồng, trong đó các điều khoản được phân loại theo loại rủi ro, trách nhiệm của các bên hoặc hành động xử lý rủi ro. Cụ thể, Moon và cộng sự [11] sử dụng 2.807 điều khoản được thu thập từ 56 bộ hồ sơ kỹ thuật xây dựng để nhận diện và phân loại mức độ rủi ro, đạt hiệu suất 93 %. Zhang và cộng sự [18] khai thác 3.411 điều khoản hợp đồng để thực hiện phân loại đa nhãn các điều khoản liên quan đến rủi ro, với hiệu suất 81 %. Tương tự, Pham và Han [6] sử dụng các tập dữ liệu có quy mô khác nhau để xác định rủi ro và trách nhiệm của các bên liên quan, qua đó hỗ trợ dự đoán các hành động xử lý rủi ro, đạt hiệu suất khoảng 90 %. Bên cạnh đó, Eken [17] sử dụng BERT để phân loại rủi ro, loại câu và trách nhiệm trong hợp đồng, đạt hiệu suất khoảng 73 %. Nhìn chung, các nghiên cứu cho thấy NLP có khả năng tự động hóa việc phân loại nội dung hợp đồng theo nhiều khía cạnh, đặc biệt là nhận diện rủi ro và trách nhiệm, qua đó hỗ trợ quá trình đánh giá và quản lý rủi ro trong hợp đồng xây dựng.

NLP cũng được ứng dụng để phân loại các nhóm nội dung chuyên môn trong hợp đồng và hồ sơ xây dựng, qua đó hỗ trợ tìm kiếm, kiểm tra tuân thủ và quản lý thông tin. Cụ thể, Hassan và Le [14] sử dụng 3.000 nội dung từ các hợp đồng thiết kế - thi công để phân loại các yêu

cầu phục vụ quá trình soạn thảo hợp đồng, đạt hiệu suất 93 %. Salama và El-Gohary [15] khai thác 330 mẫu từ các điều kiện hợp đồng và quy định để phân loại các danh mục kiểm tra tuân thủ, đạt hiệu suất 100 %, qua đó hỗ trợ tự động xác định các nội dung cần kiểm tra. Trong lĩnh vực quản lý chi phí, Deza và cộng sự [12] sử dụng 50.000 mẫu dữ liệu chi phí kết hợp với TCN để phân loại các tiêu chuẩn đo lường chi phí xây dựng, đạt hiệu suất 90 %, góp phần chuẩn hóa và tổ chức thông tin chi phí. Bên cạnh đó, Jafari và cộng sự [16] áp dụng XGBoost trên 8.943 mẫu dữ liệu từ hợp đồng và hồ sơ kỹ thuật để nhận diện các điều khoản liên quan đến yêu cầu báo cáo, đạt hiệu suất 99 %, qua đó hỗ trợ tự động hóa việc trích xuất nghĩa vụ báo cáo và dự đoán các yếu tố liên quan đến thời gian và chi phí.

Một hướng ứng dụng khác của NLP là tự động nhận diện điều khoản bị thiếu, qua đó hỗ trợ quá trình rà soát hợp đồng. Chẳng hạn, Zhou và cộng sự [10] sử dụng 1.322 hợp đồng xây dựng thực tế để phát hiện các điều khoản bị thiếu, thông qua mô hình phân loại đa nhãn kết hợp MatchPyramid-BERT, đạt hiệu suất 93 % trong phát hiện điều khoản thiếu và 83 % trong so khớp các điều khoản tương đồng. Kết quả này cho thấy NLP có tiềm năng hỗ trợ tự động sàng lọc và kiểm tra một khối lượng lớn điều khoản hợp đồng, góp phần giảm sự phụ thuộc vào quá trình rà soát thủ công.

3.2. Ứng dụng NLP trong trích xuất đặc trưng hợp đồng xây dựng

Bên cạnh đó, NLP còn được ứng dụng trong tự động trích xuất các đặc trưng ngữ nghĩa và cấu trúc, bao gồm điều khoản, thực thể, thuộc tính, quan hệ, nghĩa vụ và các yếu tố liên quan đến rủi ro từ văn bản hợp đồng (Bảng 2). Các phương pháp đa dạng, từ Rule-based NLP, Bi-GRU đến LLM kết hợp Knowledge Graph, cho phép chuyển đổi nội dung hợp đồng từ dạng văn bản tự nhiên thành dữ liệu có cấu trúc, với hiệu suất trích xuất trong nhiều nghiên cứu đạt khoảng 71-91 %. Kết quả này không chỉ hỗ trợ nhận diện rủi ro mà còn tạo nền tảng dữ liệu cho các hệ thống tự động rà soát hợp đồng, đánh giá rủi ro và hỗ trợ ra quyết định trong quản lý xây dựng.

Bảng 2. Các nghiên cứu về ứng dụng NLP trong trích xuất đặc trưng hợp đồng xây dựng.

Mô hình	Tiền huấn luyện	Mục tiêu	Hiệu suất	Số lượng data	Trích dẫn
GPT-4o	Có	Đánh giá và tóm tắt rủi ro hợp đồng	71 % (F1-score)	335	[25]
NLP	Không	Quan hệ khái niệm	90 % (F1-score)	264	[26]
NER, Phrase-Matcher, Rule-based NLP	Có	Trích xuất điều khoản rủi ro	92 % (Accuracy)	21.683	[27]
Rule-based NLP, Bi-LSTM	Không	Trích xuất điều khoản rủi ro, xếp hạng mức độ rủi ro	86 % (F1-score)	9.520	[28]
LLM	Có	Trích xuất điều khoản, đánh giá rủi ro	83 % (Hiệu suất trung bình)	155	[29]

Ứng dụng NLP trong trích xuất đặc trưng hợp đồng xây dựng chủ yếu tập trung vào việc tự động nhận diện và trích xuất các thông tin, đặc trưng và mối quan hệ quan trọng từ văn bản và điều khoản hợp đồng, qua đó giảm thời gian rà soát thủ công và hỗ trợ nhận diện rủi ro. Chẳng hạn, Lee và cộng sự [30] sử dụng mô hình Rule-based NLP để tự động trích xuất các điều khoản có tính chất bất lợi, đạt hiệu suất 81 %. trong khi đó, Qady và Kandil [26] khai thác khoảng 264 tài liệu để nhận diện các mối quan hệ giữa các khái niệm trong hợp đồng, đạt hiệu suất 90 %. Các nghiên cứu này cho thấy NLP có khả năng chuyển đổi nội dung hợp đồng từ dạng văn bản tự nhiên thành các đặc trưng có cấu trúc như điều khoản, thực thể, thuộc tính và quan hệ, qua đó tạo nền tảng cho các bước phân tích và đánh giá tiếp theo.

Một hướng nghiên cứu quan trọng khác là tự động trích xuất các nghĩa vụ, điều kiện và đặc trưng liên quan đến rủi ro từ các điều khoản hợp đồng. Cụ thể, Faraji và cộng sự [31] áp dụng phương pháp NLP dựa trên luật để xác định các nghĩa vụ trong Bộ điều khoản chung, đồng thời đánh giá mức độ bất định và rủi ro đối với mục tiêu dự án. Choi và cộng sự [27] khai thác khoảng 21.683 tài liệu, kết hợp NER, Phrase-Matcher và Rule-based NLP để tự động nhận diện các điều khoản rủi ro, đạt hiệu suất 92 %. Các nghiên cứu này cho thấy NLP có khả năng chuyển đổi nội dung câu và điều khoản hợp đồng thành các đặc trưng có cấu trúc về nghĩa vụ, trách nhiệm, điều kiện và rủi ro, qua đó hỗ trợ tự động hóa quá trình rà soát và đánh giá hợp đồng.

3.3. Ứng dụng NLP trong tạo văn bản của hợp đồng xây dựng

Nhìn chung, các nghiên cứu cho thấy NLP đang từng bước chuyển việc xử lý và tạo văn bản hợp đồng xây dựng từ phương thức đọc hiểu, tra cứu thủ công sang tự động hóa các tác vụ như tóm tắt, diễn giải, hỏi đáp, truy xuất thông tin và sinh nội dung dựa trên hợp đồng, điều khoản và các tài liệu tham chiếu. Các bài toán phổ biến bao gồm tóm tắt nội dung hợp đồng, giải thích và diễn giải điều khoản, trả lời câu hỏi liên quan đến hợp đồng, truy xuất thông tin và tự động tạo nội dung cho hồ sơ xây dựng (Bảng 3). Việc kết hợp các mô hình GPT-3.5, GPT-4, GPT-4o với RAG, Vector Database và Knowledge Graph cho thấy hiệu suất tương đối cao, với độ chính xác đạt khoảng 87-95 %. Những kết quả này cho thấy NLP có tiềm năng đáng kể trong việc giảm thời gian đọc và tra cứu hợp đồng, hỗ trợ diễn giải các điều khoản phức tạp, tự động tạo bản tóm tắt và nội dung hồ sơ, đồng thời cung cấp khả năng hỏi đáp dựa trên thông tin có căn cứ.

Bảng 3. Các nghiên cứu về ứng dụng NLP trong tạo văn bản của hợp đồng xây dựng.

Mô hình	Tiền huấn luyện	Đầu ra	Hiệu suất	Số lượng data	Trích dẫn
GPT-3.5	Có	Giải thích, diễn giải điều khoản hợp đồng xây dựng	88 % (Hiệu suất trung bình)	200	[32]
Dual-graph RAG kết hợp GPT-4o	Có	Câu trả lời được truy xuất và tổng hợp từ các điều khoản liên kết với nhau	87 % (F1 score)	93	[33]
Text-to-Text Transfer Transformer (T5)	Có	Dịch thuật và tóm tắt	92 % (F1 score)	17.709	[34]
GPT-4	Có	Bản tóm tắt nội dung hợp đồng, câu trả lời cho câu hỏi về điều khoản	93 % (Tỷ lệ hài lòng)	69.000	[35]
BERT	Có	Câu trả lời trích xuất từ văn bản	83 % (F1 score)	533	[36]
LLM	Có	Câu trả lời cho câu hỏi tri thức hợp đồng	87 % (Tỷ lệ thành công)	3.802	[37]

NLP được sử dụng để tự động tóm tắt, diễn giải và trả lời câu hỏi về nội dung hợp đồng, giúp người dùng nhanh chóng tiếp cận các thông tin và điều khoản quan trọng mà không cần đọc toàn bộ văn bản. Chẳng hạn, Gao và cộng sự [35] sử dụng GPT-4 trên các hợp đồng mẫu FIDIC để tóm tắt nội dung và trả lời câu hỏi, đạt tỷ lệ hài lòng 93 %. Saparamadu và cộng sự [32] đánh giá chatbot GPT-3.5 trên 200 cặp câu hỏi–câu trả lời về hợp đồng xây dựng, cho thấy mô hình kết hợp với Vector Database đạt hiệu suất trung bình 88 %. Lyu và cộng sự [37] xây dựng Knowledge Graph từ toàn bộ FIDIC Yellow Book 2022 và New Engineering Contract, Edition 4 để hỗ trợ truy xuất, hỏi đáp và diễn giải tri thức hợp đồng, đạt tỷ lệ thành công trên 87 %. Bên cạnh đó, Zhang và cộng sự [38] sử dụng LLM phát triển mô hình hỏi-đáp về hợp đồng dành riêng cho lĩnh vực xây dựng. Nhìn chung, các nghiên cứu cho thấy NLP có khả năng hỗ trợ hiệu quả việc tóm tắt, tra cứu và diễn giải nội dung hợp đồng, qua đó rút ngắn thời gian đọc hiểu và tiếp cận thông tin trong quá trình quản lý hợp đồng xây dựng.

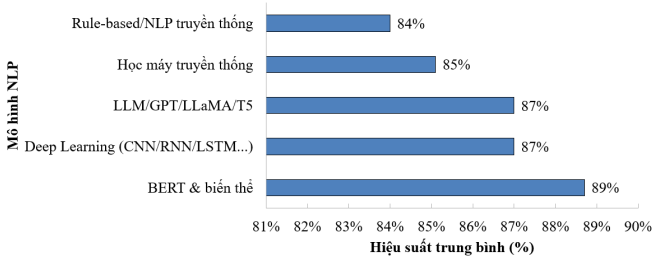
Bên cạnh đó, NLP cũng được ứng dụng để tự động tạo sinh nội dung dựa trên dữ liệu hợp đồng và hồ sơ xây dựng, chẳng hạn như bản tóm tắt, nội dung hồ sơ thầu hoặc các đoạn văn phục vụ quá trình rà soát hợp đồng. Zhao và Li [39] sử dụng LLM kết hợp RAG để tự động sinh nội dung từ các hồ sơ thầu bán cấu trúc. Trong khi đó, Zhou và cộng sự [40] khai thác 144 mẫu hồ sơ mời thầu để xây dựng các chuỗi template và module nội dung, từ đó tự động tạo nội dung cho hồ sơ mới, đạt hiệu suất lên tới 94%. Các kết quả này cho thấy NLP không chỉ hỗ trợ trích xuất và phân tích thông tin mà còn có khả năng tái tổ chức, tổng hợp và tạo sinh nội dung dựa trên các hợp đồng, điều khoản và tài liệu tham chiếu hiện có, qua đó hỗ trợ quá trình soạn thảo, chuẩn hóa và quản lý tài liệu xây dựng.

4. Các yếu tố tác động đến hiệu suất các mô hình NLP trong quản lý hợp đồng xây dựng

Nghiên cứu này đánh giá các mô hình NLP thông qua: (1) so sánh hiệu suất của các mô hình, (2) khảo sát tầm ảnh hưởng của tiền huấn luyện và quy mô dữ liệu đến hiệu suất và (3) so sánh hiệu suất mô hình đơn nhiệm và đa nhiệm theo quy mô dữ liệu.

4.1. Hiệu suất các mô hình NLP

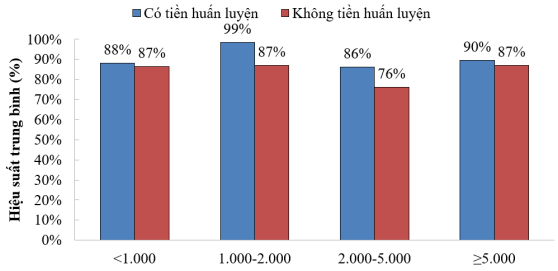
Xét về hiệu suất trung bình giữa các nhóm mô hình, Hình 8 cho thấy nhóm BERT & biến thể đạt hiệu suất cao nhất với 89 %, theo sau là các nhóm học sâu và LLM/GPT/LLaMA/T5 với 87 %. Các nhóm này vượt trội so với nhóm học máy truyền thống đạt 85 % và rule-based/NLP truyền thống với 84 %. Kết quả này cho thấy các mô hình dựa trên kiến trúc Transformer được tiền huấn luyện có ưu thế rõ rệt trong việc xử lý ngôn ngữ phức tạp và giàu ngữ cảnh của hợp đồng xây dựng so với các phương pháp truyền thống.



Hình 8. Hiệu suất trung bình của các mô hình NLP phổ biến.

4.2. Ảnh hưởng của tiền huấn luyện và quy mô dữ liệu đến hiệu suất

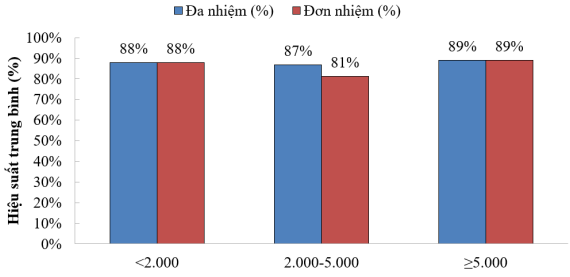
Nghiên cứu này đã điều tra hiệu suất trung bình của các mô hình NLP dưới sự tác động của tiền huấn luyện và quy mô dữ liệu, được thể hiện ở Hình 9. Ở hầu hết các nhóm quy mô dữ liệu, mô hình có tiền huấn luyện cho hiệu suất trung bình cao hơn mô hình không tiền huấn luyện. Chênh lệch rõ nhất ở nhóm dữ liệu 1.000-2.000 và 2.000-5.000 mẫu (cao hơn lần lượt 12 % và 10 %), cho thấy lợi thế của các mô hình tiền huấn luyện khi lượng dữ liệu gán nhãn chuyên ngành còn hạn chế, một tình huống rất phổ biến trong lĩnh vực hợp đồng xây dựng nơi việc thu thập và gán nhãn dữ liệu đòi hỏi nhiều thời gian và chuyên môn cao.



Hình 9. So sánh hiệu suất theo quy mô dữ liệu giữa mô hình có và không có tiền huấn luyện.

Ở nhóm dữ liệu lớn (≥5.000 mẫu), khoảng cách hiệu suất thu hẹp lại còn khoảng 3 %, cho thấy khi có đủ dữ liệu chuyên ngành, các mô hình huấn luyện từ đầu có thể tiệm cận hiệu quả như của mô hình có tiền huấn luyện. Nhìn chung, kết quả củng cố gợi ý rằng đầu tư bộ dữ liệu chuyên ngành xây dựng đủ lớn có thể mang lại hiệu quả dài hạn tốt, song việc tận dụng mô hình tiền huấn luyện vẫn là lựa chọn hiệu quả khi nguồn dữ liệu gán nhãn còn hạn chế.

4.3. So sánh mô hình đơn nhiệm và đa nhiệm theo quy mô dữ liệu



Hình 10. So sánh hiệu suất mô hình đơn nhiệm và đa nhiệm theo quy mô dữ liệu.

Nghiên cứu này cũng so sánh hiệu suất mô hình đơn nhiệm và đa nhiệm theo quy mô dữ liệu, được thể hiện ở Hình 10. Kết quả cho thấy xu hướng không đồng nhất giữa hai loại mô hình tùy theo quy mô dữ liệu. Ở nhóm dữ liệu dưới 2.000 mẫu, hiệu suất hai loại mô hình gần như tương đương với 88 % do khi thiếu dữ liệu thì các mô hình đa nhiệm không thể cải thiện việc trích xuất thông tin chia sẻ và đặc trưng chung giữa các tác vụ liên quan. Ở nhóm dữ liệu 2.000-5.000 mẫu, mô hình đa nhiệm thể hiện lợi thế rõ rệt hơn, đạt 87 % so với 81 %, phản ánh khả năng chia sẻ thông tin và học đặc trưng chung giữa các tác vụ liên quan (ví dụ đồng thời nhận diện rủi ro và bên chịu trách nhiệm), giúp giảm thiểu tình trạng thiếu dữ liệu huấn luyện cho từng tác vụ riêng lẻ. Ngược lại, ở nhóm dữ liệu lớn (≥5.000 mẫu), hiệu suất hai loại mô hình gần như tương đương với 89 % do cả hai mô hình đều có thể tối ưu hóa sâu, chuyên biệt cho một tác vụ cụ thể khi đã có đủ dữ liệu huấn luyện.

5. Những thách thức và hướng nghiên cứu trong tương lai

Kết quả tổng hợp và phân tích cho thấy lĩnh vực ứng dụng NLP vào quản lý hợp đồng xây dựng đang phát triển nhanh nhưng vẫn còn ở giai đoạn tương đối sơ khai xét trên nhiều khía cạnh: phạm vi tài liệu được khai thác, khả năng thích ứng của mô hình ngôn ngữ lớn với miền dữ liệu chuyên ngành hẹp, quy mô và chất lượng dữ liệu huấn luyện, cũng như tính đa dạng về ngôn ngữ và cấu trúc tác vụ. Từ đó, bốn nhóm thách thức chính được phân tích lần lượt trong các mục 5.1-5.4, bao gồm: (1) mất cân đối trong phạm vi tài liệu và bài toán nghiên cứu, đặc biệt ở các mảng còn ít được khai thác như tranh chấp và khiếu nại; (2) hạn chế trong khả năng thích ứng của mô hình ngôn ngữ lớn với miền dữ liệu chuyên ngành hẹp; (3) thiếu bộ dữ liệu chuẩn hóa, quy mô lớn và công khai phục vụ đánh giá mô hình; và (4) hạn chế về lệch pha ngôn ngữ và tính đa nhiệm. Trên cơ sở đó, các giải pháp được đề xuất tương ứng bao gồm: (i) mở rộng xây dựng bộ dữ liệu "hồ sơ tranh chấp/khiếu nại", (ii) xây dựng các mô hình ngôn ngữ được tiền huấn luyện trên kho dữ liệu hợp đồng xây dựng đa dạng, (iii) xây dựng và công bố các bộ dữ liệu chuẩn hóa dùng chung cho cộng đồng nghiên cứu quản lý hợp đồng xây dựng, và (iv) xây dựng và mở rộng các bộ dữ liệu hợp đồng xây dựng tiếng Việt và phát triển các kiến trúc đa nhiệm.

5.1. Mất cân đối trong phạm vi tài liệu và bài toán nghiên cứu

Như đã trình bày ở Hình 5, trong số các nhóm tài liệu/điều khoản hợp đồng được khảo sát, nhóm "hồ sơ tranh chấp/khiếu nại" chiếm tỷ lệ thấp nhất trong toàn bộ tập dữ liệu (6 %), thấp hơn nhiều so với nhóm "điều khoản hợp đồng" (46 %) hay "điều luật/văn bản pháp lý" (19 %). Đây là một nghịch lý đáng chú ý: tranh chấp và khiếu nại là giai đoạn tốn kém, kéo dài và có tác động tài chính - pháp lý lớn trong vòng đời một dự án xây dựng, nhưng lại là mảng dữ liệu ít được khai thác nhất bằng NLP. Nguyên nhân sâu xa đến từ việc hồ sơ tranh chấp thường mang tính bảo mật cao (do liên quan đến lợi ích thương mại, uy tín doanh nghiệp và các thỏa thuận không tiết lộ), khiến việc thu thập dữ liệu công khai gặp nhiều rào cản pháp lý. Hướng nghiên cứu đề xuất cho khoảng trống là hợp tác với các công ty, tổ chức xây dựng hoặc các hiệp hội nghề nghiệp để xây dựng bộ dữ liệu "hồ sơ tranh chấp/khiếu nại" được ẩn danh hóa nhưng vẫn giữ được cấu trúc ngữ nghĩa cần thiết cho huấn luyện mô hình.

5.2. Hạn chế của các mô hình ngôn ngữ lớn đã được tiền huấn luyện khi áp dụng trên miền dữ liệu hợp đồng xây dựng

Dữ liệu khảo sát cho thấy phần lớn nghiên cứu ứng dụng mô hình ngôn ngữ lớn trong quản lý hợp đồng xây dựng hiện vẫn dừng ở việc sử dụng mô hình đã được tiền huấn luyện trên kho ngữ liệu tổng quát (văn bản báo chí, hội thoại, bách khoa thư) và được tinh chỉnh trên dữ liệu xây dựng. Cách tiếp cận này tuy tiết kiệm chi phí và dễ triển khai, nhưng lại phụ thuộc nhiều vào chất lượng của cơ sở tri thức bổ sung và có thể không ổn định khi gặp các điều khoản có cấu trúc

mới lạ ngoài phạm vi tri thức đã cung cấp. Văn bản hợp đồng xây dựng có đặc điểm ngôn ngữ riêng biệt - câu dài, cấu trúc lồng ghép nhiều mệnh đề điều kiện, thuật ngữ pháp lý - kỹ thuật chuyên ngành, và mức độ mơ hồ có chủ đích trong một số điều khoản. Do đó, khi áp dụng trực tiếp một mô hình ngôn ngữ lớn đã được huấn luyện tổng quát vào bài toán chuyên ngành hẹp với bộ dữ liệu huấn luyện/tinh chỉnh nhỏ, số lượng siêu tham số khổng lồ của mô hình có thể dẫn đến hiện tượng không tận dụng hết được ưu thế biểu diễn ngữ nghĩa sâu. Để giải quyết vấn đề này, một hướng nghiên cứu quan trọng cho giai đoạn tới là xây dựng các mô hình ngôn ngữ được huấn luyện trước trên kho dữ liệu hợp đồng xây dựng đa dạng - bao gồm các mẫu hợp đồng chuẩn quốc tế, quy chuẩn kỹ thuật, hồ sơ mời thầu và án lệ tranh chấp - trước khi tinh chỉnh cho từng bài toán cụ thể (nhận diện rủi ro, trích xuất nghĩa vụ, soạn thảo tự động).

5.3. Thiếu bộ dữ liệu chuẩn hóa, quy mô lớn và công khai phục vụ đào tạo và đánh giá mô hình

Rào cản tiếp cận dữ liệu hợp đồng xây dựng thực tế vẫn còn rất lớn, chủ yếu do tính bảo mật thương mại và sự phân mảnh giữa các doanh nghiệp, quốc gia và loại hình dự án. Hệ quả trực tiếp là: (1) khó so sánh khách quan hiệu suất giữa các phương pháp khác nhau khi mỗi nghiên cứu dùng một bộ dữ liệu riêng với quy mô, ngôn ngữ và tiêu chí gán nhãn khác nhau; và (2) các mô hình được huấn luyện trên dữ liệu nhỏ, đặc thù cho một loại hợp đồng hoặc một quốc gia cụ thể có nguy cơ không được tổng quát hóa tốt khi áp dụng cho loại hợp đồng hoặc khu vực pháp lý khác.

Để giải quyết khoảng trống này, hướng nghiên cứu tương lai cần tập trung vào việc xây dựng và công bố các bộ dữ liệu chuẩn hóa dùng chung cho cộng đồng nghiên cứu quản lý hợp đồng xây dựng. Một bộ dữ liệu chuẩn hóa như vậy cần đảm bảo: quy mô đủ lớn để huấn luyện và đánh giá công bằng cả mô hình học máy truyền thống lẫn mô hình sâu và LLM; đa dạng về loại hợp đồng và loại tác vụ; có quy trình gán nhãn được chuẩn hóa và kiểm định bởi chuyên gia trong ngành để đảm bảo chất lượng nhãn; đồng thời cần có cơ chế ẩn danh hóa hoặc tổng hợp hóa phù hợp để dung hòa giữa nhu cầu chia sẻ dữ liệu phục vụ nghiên cứu và yêu cầu bảo mật thương mại của các bên liên quan.

5.4. Hạn chế về lệch pha ngôn ngữ, tính đa nhiệm

Một hạn chế đáng chú ý khác là sự lệch pha về ngôn ngữ: đa số các nghiên cứu được khảo sát sử dụng dữ liệu tiếng Anh, trong khi các bộ dữ liệu hợp đồng xây dựng bằng tiếng Việt - hoặc các ngôn ngữ khác có nguồn tài nguyên NLP còn hạn chế. Điều này tạo ra một khoảng trống lớn đối với các quốc gia đang phát triển, nơi phần lớn hợp đồng xây dựng nội địa được soạn thảo bằng ngôn ngữ bản địa và chịu sự chi phối của hệ thống pháp luật riêng, khiến các mô hình tiền huấn luyện trên dữ liệu tiếng Anh khó có thể áp dụng trực tiếp mà không qua một bước phiên dịch và tinh chỉnh đáng kể. Bên cạnh đó, kết quả phân tích ở Mục 4.3 cho thấy mô hình đa nhiệm có ưu thế rõ rệt ở quy mô dữ liệu vừa phải nhờ khả năng chia sẻ đặc trưng giữa các tác vụ liên quan, nhưng phần lớn nghiên cứu hiện nay vẫn tập trung xây dựng mô hình đơn nhiệm chuyên biệt cho một bài toán cụ thể (ví dụ chỉ phân loại rủi ro hoặc chỉ trích xuất điều khoản), dẫn đến sự phân mảnh trong kiến trúc hệ thống và khó tích hợp thành một quy trình xử lý hợp đồng toàn diện, từ đầu đến cuối. Do đó, các hướng nghiên cứu tương lai cần được mở rộng: (1) xây dựng và mở rộng các bộ dữ liệu hợp đồng xây dựng tiếng Việt, kết hợp nghiên cứu các kỹ thuật chuyển giao tri thức liên ngôn ngữ để tận dụng các mô hình đã huấn luyện trên dữ liệu tiếng Anh cho bối cảnh tiếng Việt; và (2) phát triển các kiến trúc đa nhiệm thống nhất có khả năng đồng thời thực hiện nhận diện rủi ro, trích xuất điều khoản, phát hiện bất thường và hỗ trợ soạn thảo trong một quy trình liền mạch, thay vì các mô hình rời rạc cho từng tác vụ.

6. Kết luận

Nghiên cứu này đã tổng hợp và phân tích 124 nghiên cứu về ứng dụng NLP trong quản lý hợp đồng xây dựng, tập trung phân tích xu hướng nghiên cứu trong quản lý hợp đồng xây dựng thông qua việc phân loại theo nhóm tác vụ, loại tài liệu và phương pháp, đồng thời đánh giá ảnh hưởng của việc sử dụng mô hình tiền huấn luyện, quy mô dữ liệu huấn luyện và kiến trúc đơn nhiệm/đa nhiệm đến hiệu suất mô hình. Kết quả cho thấy phân loại văn bản là nhóm nhiệm vụ phổ biến nhất (60 %), tiếp theo là trích xuất thông tin (29 %) và tạo sinh văn bản (11 %). Trong số các tài liệu hợp đồng, “điều khoản hợp đồng” chiếm tỷ trọng cao nhất (46 %). Các mô hình học máy truyền thống chiếm tỷ lệ cao nhất (26 %), trong khi BERT và các biến thể đạt hiệu suất trung bình cao nhất (89 %). Phân tích theo quy mô dữ liệu cho thấy các mô hình được tiền huấn luyện có lợi thế rõ rệt khi dữ liệu gán nhãn hạn chế, nhưng khoảng cách này thu hẹp khi quy mô dữ liệu tăng. Tương tự, mô hình đa nhiệm có xu hướng đạt hiệu suất tốt hơn ở quy mô dữ liệu trung bình, trong khi mô hình đơn nhiệm trở nên cạnh tranh hơn khi có nhiều dữ liệu.

Về hạn chế và hướng nghiên cứu, các nghiên cứu hiện nay còn mất cân đối về phạm vi tài liệu và nhiệm vụ, đặc biệt thiếu các nghiên cứu về tranh chấp và khiếu nại; khả năng thích ứng của mô hình ngôn ngữ lớn với miền hợp đồng xây dựng còn hạn chế; thiếu các bộ dữ liệu chuẩn hóa, quy mô lớn và công khai; đồng thời còn hạn chế về dữ liệu tiếng Việt và kiến trúc đa nhiệm. Do đó, các nghiên cứu tương lai cần mở rộng xây dựng bộ dữ liệu về tranh chấp và khiếu nại; phát triển các mô hình ngôn ngữ được huấn luyện trước trên dữ liệu hợp đồng xây dựng chuyên ngành; xây dựng, phát triển và công bố các bộ dữ liệu hợp đồng xây dựng tiếng Việt cũng như các kiến trúc đa nhiệm nhằm nâng cao khả năng ứng dụng NLP trong quản lý hợp đồng xây dựng. Với sự phát triển nhanh chóng và mức độ tích hợp ngày càng gia tăng của NLP trong lĩnh vực hợp đồng xây dựng, nghiên cứu này cung cấp những hiểu biết có giá trị về các chiến lược ứng dụng NLP phù hợp trong lĩnh vực này, đồng thời định hướng cho các nghiên cứu trong tương lai.

Lời cảm ơn

Nhóm tác giả trân trọng cảm ơn Trường Đại học Công nghệ Kỹ thuật TP. HCM đã tài trợ theo hợp đồng đề tài NCKH SV mã số: 345/HĐ-KHCN-SV ký ngày 27/01/2026

Tài liệu tham khảo

[1]. S.-O. Cheung and H. C. H. Suen, "A multi-attribute utility model for dispute resolution strategy selection," Constr. Manag. Econ., vol. 20, no. 7, pp. 557–568, Oct. 2002, doi: 10.1080/01446190210157568.

[2]. M. Abdelsayed and R. Moffat, "Design on better outcomes, getting to the CRUX of time and cost overruns on capital projects," ReNew Canada.

[3]. The Construction Broadsheet, "Report: Value of global construction claims reach $73B," 2021, [Online]. Available: https://theconstructionbroadsheet.com/report-value-of-global-construction-claims-reach-b-p544-117.htm

[4]. R. Khalef and I. H. El-adaway, "Automated Identification of Substantial Changes in Construction Projects of Airport Improvement Program: Machine Learning and Natural Language Processing Comparative Analysis," J. Manag. Eng., vol. 37, no. 6, p. 04021062, 2021, doi: 10.1061/(ASCE)ME.1943-5479.0000959.

[5]. I. Dikmen, G. Eken, H. Erol, and M. T. Birgonul, "Automated construction contract analysis for risk and responsibility assessment using natural language processing and machine learning," Comput. Ind., vol. 166, p. 104251, 2025, doi: 10.1016/j.compind.2025.104251.

[6]. H. T. T. L. Pham and S. Han, "Natural Language Processing with Multitask Classification for Semantic Prediction of Risk-Handling Actions in Construction Contracts," J. Comput. Civ. Eng., vol. 37, no. 6, 2023, doi: 10.1061/jccee5.cpeng-5218.

[7]. M. Soy, T. Zhao, H. Hegazy, J. Zhang, and E. Elwakil, "Advancement of Construction Contract Analysis with Natural Language Processing and Large Language Models: A Literature Review," 2026, doi: 10.22260/crc-csce-2025/0104.

[8]. M. S. Uddin, S. D. Datta, A. Abdelaty, and M. Babar, "Application of Artificial Intelligence in Construction Contract Management: A Systematic Review," Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, vol. 18, no. 3, p. 03126002, 2026. doi: 10.1061/jladah.ladr-1536.

[9]. E. O. Aktas, B. Un, and O. Genc, "Applications of Artificial Intelligence in Construction Contract Management," 5th International Conference on Modern and Advanced Research, Turkey, pp. 357–364, 2026.

[10]. H. Zhou, B. Gao, S. Tang, B. Li, and S. Wang, "Intelligent detection on construction project contract missing clauses based on deep learning and NLP," Eng. Constr. Archit. Manag., vol. 32, no. 3, pp. 1546–1580, 2025, doi: 10.1108/ECAM-02-2023-0172.

[11]. S. Moon, S. Chi, and S. B. Im, "Automated detection of contractual risk clauses from construction specifications using bidirectional encoder representations from transformers (BERT)," Autom. Constr., vol. 142, p. 104465, 2022, doi: 10.1016/j.autcon.2022.104465.

[12]. J. I. Deza, H. Ihshaish, and L. Mahdjoubi, "A Machine Learning Approach to Classifying Construction Cost Documents into the International Construction Measurement Standard," 2022, doi: http://arxiv.org/abs/2211.07705.

[13]. A. Erfani and H. Khanjar, "Large Language Models for Construction Risk Classification: A Comparative Study," Buildings, vol. 15, no. 18, 2025, doi: 10.3390/buildings15183379.

[14]. F. ul Hassan and T. Le, "Computer-assisted separation of design-build contract requirements to support subcontract drafting," Autom. Constr., vol. 122, p. 103479, 2021, doi: 10.1016/j.autcon.2020.103479.

[15]. D. M. Salama and N. M. El-Gohary, "Semantic Text Classification for Supporting Automated Compliance Checking in Construction," J. Comput. Civ. Eng., vol. 30, no. 1, p. 04014106, 2016, doi: 10.1061/(asce)cp.1943-5487.0000301.

[16]. P. Jafari, M. Al Hattab, E. Mohamed, and S. Abourizk, "Automated extraction and time-cost prediction of contractual reporting requirements in construction using natural language processing and simulation," Appl. Sci., vol. 11, no. 13, 2021, doi: 10.3390/app11136188.

[17]. G. Eken, "Using natural language processing for automated construction contract review during risk assessment at the bidding stage," Ph.D. - Doctoral Program, Middle East Technical University, pp. 1–154, 2022. [Online]. Available: https://open.metu.edu.tr/handle/11511/101793

[18]. L. Zhang, Y. Ning, and S. Zhou, "Automated multi-label classification of risk clauses in construction contracts using GPT-driven data augmentation," Autom. Constr., vol. 181, p. 106599, 2026, doi: 10.1016/j.autcon.2025.106599.

[19]. P. Waditwar, "Agentic AI in Contract Analytics Harnessing Machine Learning for Risk Assessment and Compliance in Government Procurement Contracts," Open J. Bus. Manag., vol. 13, no. 05, pp. 3385–3395, 2025, doi: 10.4236/ojbm.2025.135179.

[20]. G. Eken, I. Dikmen, and M. T. Birgonul, "Using NLP for Automated Contract Review and Risk Assessment," Proceedings ofthe 33rd European Safety and ReliabilityConference (ESREL 2023), pp. 2566–2567, 2023. doi: 10.3850/978-981-18-8071-1_p030-cd.

[21]. R. Bashir, M. Salamak, and S. Ahmed, "Dispute Management in Engineering Contracts Using Artificial Intelligence," Int. J. BIM Eng. Sci., vol. 11, no. 2, 2025, doi: 10.54216/ijbes.110202.

[22]. A. Demircan and L. O. Uğur, "Analysis of Design-Build Contracts Utilising Machine Learning on the Axis of Time, Cost and Quality," Düzce Üniversitesi Bilim ve Teknol. Derg., vol. 13, no. 4, pp. 1455–1475, 2025, doi: 10.29130/dubited.1664225.

[23]. R. Balaji, K. S. Sivakumar, M. Jagannathan, and V. S. Kumar Delhi, "Automatic Assessment of Leanness of Construction Contract Provisions: A Proof-of-Concept," Proceedings of the International Symposium on Automation and Robotics in Construction, pp. 822–829, 2024. doi: 10.22260/ISARC2024/0107.

[24]. A. Demircan and L. O. Uğur, "Transformer-Based NLP for Construction Contract Clause Classification: Implications for Sustainable Construction Project Governance," Sustain., vol. 18, no. 13, 2026, doi: 10.3390/su18136788.

[25]. C. Zheng, S. Wong, X. Su, Y. Tang, A. Nawaz, and M. Kassem, "Automating construction contract review using knowledge graph-enhanced large language models," Autom. Constr., vol. 175, p. 106179, 2025, doi: 10.1016/j.autcon.2025.106179.

[26]. M. Al Qady and A. Kandil, "Concept Relation Extraction from Construction Documents Using Natural Language Processing," J. Constr. Eng. Manag., vol. 136, no. 3, pp. 294–302, 2010, doi: 10.1061/(asce)co.1943-7862.0000131.

[27]. S. J. Choi, S. W. Choi, J. H. Kim, and E. B. Lee, "Ai and text-mining applications for analyzing contractor's risk in invitation to bid (ITB) and contracts for engineering procurement and construction (EPC) projects," Energies, vol. 14, no. 15, 2021, doi: 10.3390/en14154632.

[28]. S. W. Choi and E. B. Lee, "Contractor's Risk Analysis of Engineering Procurement and Construction (EPC) Contracts Using Ontological Semantic Model and Bi-Long Short-Term Memory (LSTM) Technology," Sustain., vol. 14, no. 11, 2022, doi: 10.3390/su14116938.

[29]. E. W. Kim, Y. J. Shin, K. J. Kim, and S. Kwon, "Development of an Automated Construction Contract Review Framework Using Large Language Model and Domain Knowledge," Buildings, vol. 15, no. 6, 2025, doi: 10.3390/buildings15060923.

[30]. J. Lee, J.-S. Yi, and J. Son, "Development of Automatic-Extraction Model of Poisonous Clauses in International Construction Contracts Using Rule-Based NLP," J. Comput. Civ. Eng., vol. 33, no. 3, 2019, doi: 10.1061/(asce)cp.1943-5487.0000807.

[31]. A. Faraji, M. Rashidi, and S. Perera, "Text Mining Risk Assessment–Based Model to Conduct Uncertainty Analysis of the General Conditions of Contract in Housing Construction Projects: Case Study of the NSW GC21," J. Archit. Eng., vol. 27, no. 3, 2021, doi: 10.1061/(asce)ae.1943-5568.0000489.

[32]. P. V. I. N. Saparamadu et al., "Optimising Contract Interpretations with Large Language Models: A Comparative Evaluation of a Vector Database-Powered Chatbot vs. ChatGPT," Buildings, vol. 15, no. 7, 2025, doi: 10.3390/buildings15071144.

[33]. Z. Zhang, Y., Wang, X., Hu, M., & Zhang, "BifrostRAG Bridging Dual Knowledge Graphs for Multi-Hop Question Answering in Construction Safety," arXiv e-prints, arXiv-2507, 2025.

[34]. G. Rejithkumar, P. R. Anish, P. Sonar, and S. Ghaisas, "Automated Extraction of Compliance Elements in Software Engineering Contracts Using Natural Language Generation," Proceedings - 2024 ACM/IEEE International Workshop on NL-Based Software Engineering, NLBSE 2024, pp. 69–72, 2024. doi: 10.1145/3643787.3648030.

[35]. Y. Gao, Y. Gan, Y. Chen, and Y. Chen, "Application of large language models to intelligently analyze long construction contract texts," Constr. Manag. Econ., vol. 43, no. 3, pp. 226–242, 2025, doi: 10.1080/01446193.2024.2415676.

[36]. Y. Kim, S. Bang, J. Sohn, and H. Kim, "Question answering method for infrastructure damage information retrieval from textual data using bidirectional encoder representations from transformers," Autom. Constr., vol. 134, 2022, doi: 10.1016/j.autcon.2021.104061.

[37]. Y. Fu, T. Lyu, Y. Chen, and Z. Lv, "Q&A system for international construction contracts driven by large language model and knowledge graph," Eng. Manag., vol. 13, no. 1, pp. 124–145, 2026, doi: 10.1007/s42524-026-4237-0.

[38]. M. Zhang, C. Xu, Y. Gan, Y. Wang, Y. Fu, and Y. Chen, "Automating construction contract question answering using large language model and fine-tuning," Expert Syst. Appl., vol. 297, 2026, doi: 10.1016/j.eswa.2025.129493.

[39]. D. Zhao, Y., & Li, "A Large Language Model-Based Framework for Semi-Structured Tender Document Retrieval-Augmented Generation," arXiv preprint arXiv:2410.09077, 2024.

[40]. T. T. Zhou, X. D. Gao, and G. Y. Wei, "Construction Method of Tender Document Based on Case-based Reasoning," Int. J. Comput. Commun. Control, vol. 16, no. 3, pp. 1–20, 2021, doi: 10.15837/ijccc.2021.3.4170.