Development and Application of an Intelligent Moot Court Trial Platform Based on Machine Learning and Natural Language Processing Technology

Jia Wang, Xingyu Yuan, Yuqing Zhang, Pinxiao Guan(✉), Hui Zeng, and Zeyu Wang

Hebei University of Economics and Business, Xinhua District, Shijiazhuang, Hebei, China
1320508450@qq.com

Abstract. The continuous penetration of AI technology into the legal industry has changed the structure of the demand for legal talents, and the demand for complex legal talents in the context of AI has been further expanded, but the combination of legal education and new technology is not close enough today. We should make full use of the technical advantages of AI and combine the characteristics of traditional legal education to find a suitable path for future legal higher education reform. As the current legal education is characterized by a single research and teaching method and insufficient training of students' effective output, it is extremely important to study the use of AI technology to simulate trial teaching to cultivate practical "AI + Law" talents.

Keywords: AI Law · Judicial Trial · Machine Learning · Natural Language Processing

1 Introduction

To adapt to the general trend of promoting economic and social development and improving social governance with Big Data and AI, many universities in China are now actively responding to the Development Plan of the New Generation of Artificial Intelligence to open the discipline of AI Law, although they lack experience in training talents for this discipline [1]. As a teaching method that takes into account both theoretical teaching and practical training, the mock trials can better solve the problem of the traditional legal education model that only emphasizes knowledge transfer and academic research but neglects professional thinking training and capacity development, thus promoting the goal of "Digital Rule of Law, Smart Justice".

2 "AI + Trial" Implementation Path Planning

2.1 Function and Principle of Implementation

Real name certification, real life mock trial

Registration is required for the first login. The registration function scans and stores

D. Kumar et al. (Eds.): IEIT 2023, AHSSEH 10, pp. 558–564, 2023.
https://doi.org/10.2991/978-94-6463-230-9_67

identity information through OCR and PYTHON SQLite technologies verifies basic information such as mobile phone number and email address based on regular expressions and stores it in the database, stores images such as avatars as absolute paths in the server (or through Blob binary data types), and stores account IDs by backend code through. The account ID is generated by the backend code using random numbers and is used as the primary key. For example, Fig. 1. Three main principles of real-name login.

Select your own character role

After the user selects a case and successfully matches other users, the system randomly generates a queue of players and selects roles in order (players can voluntarily negotiate to swap the order between them) [1], and some roles will be played by the system AI as NPCs to facilitate the trial process. Once a role is selected, the system retrieves the corresponding information in SQLite using the selected case and role as keywords and presents it to the user. The images of some of the non-player-controlled roles (NPCs) are modeled as faces through deep learning simulation to help the user understand and grasp the case more visually. For example, Fig. 1. Three main principles of select cases and roles.

Promote relevant legal information

The legal knowledge reasoning is carried out through techniques such as legal knowledge mapping, where legal knowledge points such as certain legal provisions, legal documents and evidential materials are strung together to form a conceptual framework on which the mock trial is driven. The case itself and the legal provisions involved as the trial progresses are pushed to the interface in real-time [2].

Online matchmaking chat

The instant messaging technology provided by LeanCloud enables ad hoc conversations to be held during the mock trial to make statements and questions to make statements and questions during the mock trial. Temporary group chats can also be set up by users in different roles to share information and ideas [3].

Selecting valid evidence

When the user collects evidence, the system backend will scan the text material through OCR text recognition and then analyze and process the image file to obtain text and layout information and display it next to the corresponding original text or in the form of annotations for the player to observe the original evidence images and collect the rest of the non-textual information [4]. Some parts of the different pieces of evidence are accompanied by textual descriptions to facilitate the trial process. Evidence that the player considers valid can be flagged for subsequent review, use or comparison with other evidence.

Automatic generation of legal documents

The automatic legal document generation function automatically crawls public legal documents through PYTHON crawler technology and parses and refines legal knowledge based on Natural Language Understanding (NLP) and Deep Learning (DL) technologies, establishes data fusion rules through legal data fusion technology, and then realizes the

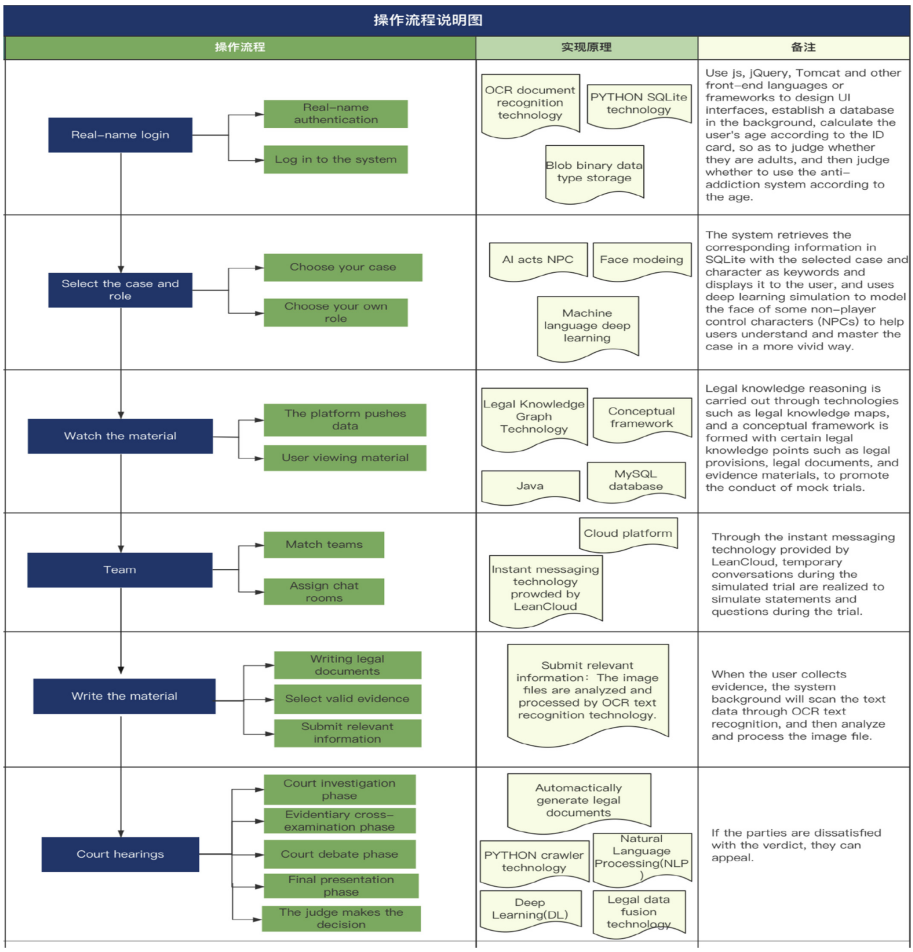

Fig. 1. Diagram of the operating process

fusion of structured, semi-structured and unstructured multi-source heterogeneous data based on semantics to build a relationship model [5]. The data is then identified and established based on legal semantics, solving the problems of data integrity, consistency and relevance, and forming a highly integrated data resource map. The legal knowledge map is constructed through the knowledge domain visualization technology, which connects the legal knowledge points in legal regulations, legal documents, evidence materials, and other legal materials with certain legal logic to form a conceptual framework. The final use of specific elements triggers words on the event elements through the template library for analysis, editing, and other operations to generate legal documents [2].

2.2 Operating Procedures

Step 1: First, users can enter and log in the system with real-name authentication. After entering the game interface, they can select the cases they are interested in through case type page sorting or searching keywords directly in the search bar.

```
function authenticateUser(username, password)
function loginUser(username, password) {
    if (authenticateUser(username, password)) {
        return getUserInfo(username);
    } else {
        return { error: "Authentication failed" };
    }
}
function getCaseList(pageNumber, pageSize, keyword) {
searchCases(keyword)
}
function searchCases(keyword) {
    const caseList = getCaseList(1, 10, keyword);
    renderCaseList(caseList);
}
```

Step 2: After selecting the role, users will be assigned to a chat room, who can write the indictment, defense and other documents and collect relevant evidence.

```
function joinChatRoom(caseId, userId) {
    const relatedUsers = getRelatedUsers(caseId, userId).
    joinRoomAndShowChat(relatedUsers).
}
// The user writes an indictment and a reply
function writeIndictment(caseId, userId, content) {
    saveIndictment(caseId, userId, content).
}
function writeResponse(caseId, userId, content) {
    saveResponse(caseId, userId, content).
}
// The user uploads evidence
function uploadEvidence(caseId, userId, evidence) {
    saveEvidence(caseId, userId, evidence).
}
function searchRelatedMaterial(keyword) {
    const material = searchMaterial(keyword).
// Display the search results screen render Search Result(material).
}
```

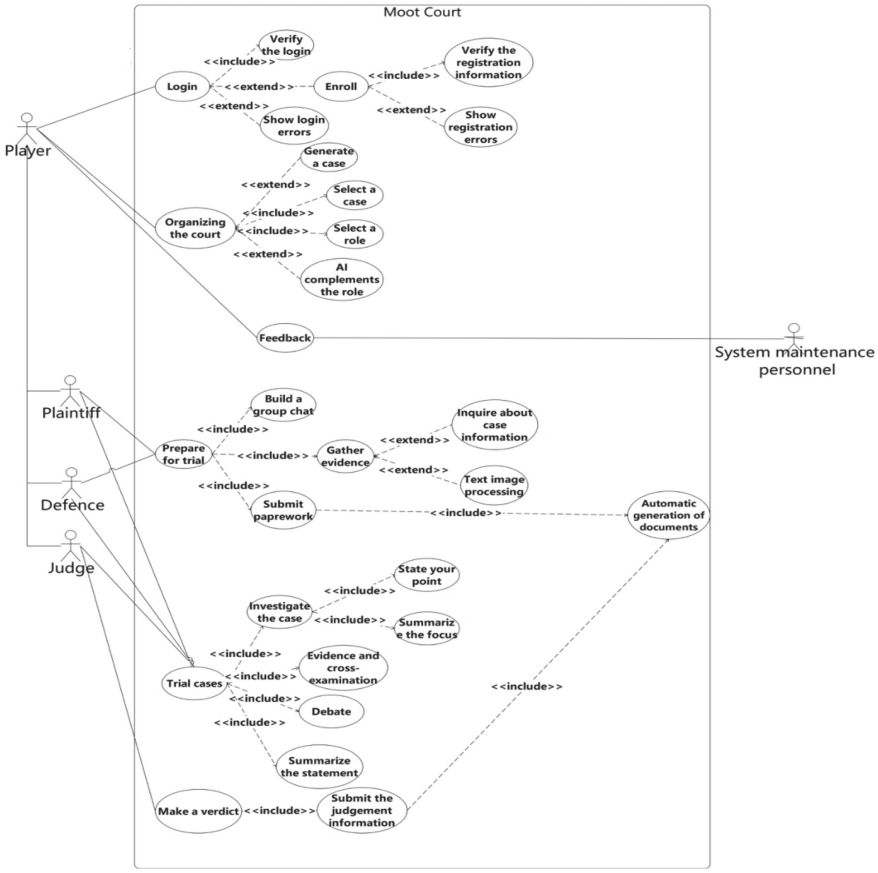

Fig. 2. Schematic

Step 3: After completing the above, the user enters the court hearing stage, which is divided into five main stages, namely the court investigation stage, the evidence presentation and cross-examination stage, the court argument stage, the final statement stage and the judge's discussion followed by the delivery of the judgement and the judgment stage. For example, Fig. 2 shows the specific operation process.

```
// Enter the Investigation Phase,
function enterInvestigationPhase(caseId) {
    renderInvestigationPhase(caseId).
}
// Enter the Evidence Phase, querying the database for information about the indictment, pleadings and evidence related to the case
function enterEvidencePhase(caseId) {
    const indictment = getIndictment(caseId).
    const response = getResponse(caseId).
    const evidence = getEvidence(caseId).
    renderEvidencePhase(indictment, response, evidence).
}
// Enter the courtroom argument phase
function enterDebatePhase(caseId) {
    renderDebatePhase(caseId).
}
// Enter the final statement phase
function enterFinalStatementPhase(caseId) {
    renderFinalStatementPhase(caseId).
}
// Enter the judge discussion phase and query the database for the judge of the case
function enterJudgeDiscussionPhase(caseId)
    deliverVerdict(verdict).
```

3 Prospects and Challenges

Under the impact of the wave of artificial intelligence, the digital characteristics of legal practice work will only become more obvious. The society should actively face this situation, but from the present point of view, the computer level of law school students is still at the ordinary secondary level, unable to cultivate legal compound practical talents for the society [2]. Followed by research, university law professional students for law practice demand greatly and very urgent, this platform belongs to the wisdom of the court in the future, combining AI technology and legal education, by entertaining the franco-prussian and legal study, overcome the disadvantages of the current legal education theory and practice.

However, there are risks associated with the platform that cannot be ignored. The artificial intelligence technology invoked by the platform is prone to legal risks. The increasing anthropomorphism and intelligence of today's artificial intelligence products bring with it the challenge of artificial intelligence on data security. Without the support of strong technical conditions and professionals, the risk of data leakage and theft will persist. Data is also essential to the normal operation of the platform, but it is extremely difficult to collect. And the problem of lack of data still exists [3].

4 Conclusions

The platform uses machine learning, natural language processing and other technologies to simulate the trial process, restoring the whole process of handling cases from case follow-up to court trials, exercising students' practical skills through evidence collection, document writing and court arguments, etc. In addition to this, it also provides feedback on the handling of the specific process of the case, playing the role of a teacher in real teaching, allowing students to recognize their shortcomings in each case and thus continuously improve their professional competence. The platform provides a platform for students to practice with the help of artificial intelligence technology and to develop "artificial intelligence + law" talents [5]. The platform is to provide a platform for practical practice with the help of artificial intelligence technology and to nurture "AI + Law" talents.

Acknowledgment. Provincial Entrepreneurship Innovation Training Project for University Students, e-Case Delivering Law - Immersive Role-playing Moot Court Platform, 202311832071X.

References

1. Shi Ya, Dai Xianhua, Zhao Liang, Deng Shanxi. (2022) Exploring the reform of experimental teaching of law in local universities in the context of artificial intelligence [J]. Laboratory Research and Exploration, 41(05):234–237+243.
2. Wei B, Xiang Kun, Sun Changlong, Feng Jun, Zhang Yating, Zhu Xinli, Zhou Jianghong, Zhai Yinsheng, Wu Fei. (2022) A whole-process intelligent trial system for smart courts (in English) [J]. Frontiers of Information Technology & Electronic Engineering, 23(02): 186-207.
3. Wei B. (2022) Legal artificial intelligence: scientific connotation, evolutionary logic and trend foresight [J]. Journal of Zhejiang University (Humanities and Social Sciences Edition), 52(07):49-67.
4. Zhou M, Duan N, Liu Shujie et al. (2022) Recent advances in neural natural language processing - models, training, and inference [J]. Engineering, 6(03):155-188.
5. Zuo Weimin. (2018) Some thoughts on the prospect of using legal artificial intelligence in China [J]. Tsinghua Jurisprudence, 12(02):108-124.