International Journal of Education and Humanities
ISSN: 2770-6702 | Vol. 24, No. 1, 2026

Development and Teaching Practice of Graduate Course Resources for "AI + Finance": A Case Study of Generative Artificial Intelligence and Financial Applications

Dongjie Lin*

School of Public Finance and Taxation, Central University of Finance and Economics, Beijing 102206, China
* **Corresponding author email:** rayldj@cufe.edu.cn

Abstract: This paper discusses the development of course resources and teaching practice for the graduate course Generative Artificial Intelligence and Financial Applications. Built on earlier work in financial data analysis and Python-based teaching, the course responds to the growing need for graduate students to use generative artificial intelligence in annual report reading, policy analysis, financial statement interpretation, financial news assessment, and research report writing. It develops a resource system composed of conceptual lectures, case materials, experimental task packages, reading assignments, process portfolios, verification logs, and assessment rubrics. The course organizes model use around authentic financial materials and disciplinary problems, with emphasis on source verification, evidence-based interpretation, and normative reflection. It guides students to form financial analysis paths that can be traced, explained, and questioned under model-assisted conditions. The paper focuses on the course resource framework, teaching process design, key risk responses, and future evaluation plans. Quantitative assessment of teaching effects will be conducted in later teaching cycles.

Keywords: AI + Finance, Generative Artificial Intelligence, Graduate Education.

1. Introduction

Generative artificial intelligence, represented by large language models, is changing the basic ways in which people search for knowledge, generate text, write code, and organize materials. Model development in recent years has brought natural language interaction, long-text understanding, table interpretation, code generation, and multi-turn dialogue into ordinary learning settings. For university teaching, the impact is not limited to whether students can complete an assignment more quickly. The chain of knowledge acquisition, material processing, argument formation, and learning assessment is also being reorganized. UNESCO's guidance on the use of generative AI in education and research, together with its AI competency framework for teachers, has brought model capability, ethical boundaries, and teaching responsibility into discussions of educational digitalization [1-2]. In graduate education for finance-related disciplines, this change is already visible. Students use models to search policy backgrounds, summarize corporate annual reports, explain statement items, generate data-processing code, prepare research report outlines, and revise the language of course papers. If teachers continue to treat these practices as extracurricular tool use, they will find it difficult to understand students' actual learning processes or judge whether disciplinary competence has been trained.

For finance-related graduate students, learning generative artificial intelligence is not mainly a matter of mastering the interface of a particular model. The central issue is whether model assistance can be turned into verifiable professional analysis. Financial materials are often long, heterogeneous, and full of conditions. A corporate announcement may need to be read together with audit opinions, regulatory inquiry letters, industry conditions, and management discussion. The effect of a tax policy depends on the taxpayer, income category, regional scope, effective date, and supporting rules. The interpretation of a financial ratio must return to the formula, accounting definition, industry context, and original data. A model may quickly produce fluent text, yet it may overlook conditions that are decisive for financial judgment. It may also mix together policy clauses, accounting concepts, or market sentiments that look similar on the surface. Finance education therefore cannot stop at asking whether students can use a model. It also has to train them to define tasks, choose sources, verify outputs, explain evidence, and take responsibility for conclusions.

The graduate course Generative Artificial Intelligence and Financial Applications was developed against this background. It is designed for graduate students in public finance, taxation, finance, accounting, asset appraisal, digital economy, and related fields. Its foundation comes from earlier finance data courses, Python teaching, MOOC development, textbook development, and financial case construction. Earlier teaching placed more weight on programming, data cleaning, deep learning models, and financial modeling tasks. After the course was renamed, the teaching focus moved toward financial text processing, material organization, output verification, and risk governance. This shift is not a simple replacement of prior data courses. It places generative artificial intelligence within the professional learning process of finance on the basis of existing data-analysis training. Taking this course as an example, this paper discusses how course resources can be reorganized after generative artificial intelligence enters course papers, case analysis, experimental tasks, and research report writing; how classroom tasks can be designed; and how assessment can move from final texts to traceable learning processes.

This paper is a study of course construction. Its basis includes literature from the past three years on generative AI in education, financial applications, and governance risks, as well as the teaching foundations already formed in course practice and project application materials. It does not report questionnaire results, grade changes, or student satisfaction data that have not yet been collected, nor does it treat scattered classroom impressions as empirical conclusions. The paper explains how course resources and teaching processes are designed, which teaching links can be recorded and assessed, and what kinds of evidence later empirical research should collect. This approach helps locate course reform at the levels of resources, tasks, assessment, and norms, rather than turning course construction in generative artificial intelligence into a general introduction to technology.

2. Research Basis and Course Design Logic

2.1. From AI Literacy to the Reconstruction of Finance Courses

Research on generative artificial intelligence in higher education has moved from whether it can be used to how it can be embedded in courses. Systematic reviews show that generative models can support lesson preparation, automated feedback, literature organization, code assistance, and personalized learning, while also creating problems related to unreliable sources, assessment validity, academic integrity, and student dependence [3-7]. Lo's rapid review of ChatGPT's impact on education offers an important judgment for course assessment: if a course evaluates only the final text, teachers can hardly tell whether students have gone through a real learning process [7]. This judgment is especially relevant to finance courses. The final form of a financial report may be a paragraph of explanation, a table, or an analytical text, yet behind it lie sources, calculations, institutional rules, and professional judgments. If students submit only a finished draft, teachers can hardly distinguish which views come from public documents, which come from model suggestions, which have been manually verified, and which are merely plausible guesses.

Research on university assessment policy also suggests that a course should not focus narrowly on whether a work is original. Luo's critical review of assessment policies on generative artificial intelligence points out that many policies stress the originality of student work, while giving less attention to whether assignments still measure the abilities that courses intend to cultivate [8]. Moorhouse, Yeo, and Wan, in their analysis of assessment guidelines from top-ranked universities, emphasize that tool boundaries, assessment validity, and academic integrity all need to be included in course design [9]. For Generative Artificial Intelligence and Financial Applications, the core abilities to be assessed include raising financial questions, selecting materials, defining model tasks, verifying the basis of model outputs, and producing explainable professional interpretations. The course therefore cannot simply prohibit models, nor can it allow students to treat model output as a conclusion. A more workable approach is to include model use in a learning process that can be recorded, questioned, and assessed.

This design corresponds to the basic requirements of graduate education. Graduate study emphasizes problem awareness, judgment of evidence, and revision of viewpoints. The links most easily affected by generative artificial intelligence are precisely material organization, first-draft production, and language revision. A model may lower the threshold for getting started, but it cannot replace students' judgment about research questions, evidential weight, and the boundaries of conclusions. Financial problems make this difference clear. Risk disclosure in an annual report may be formulaic, or it may indicate changes in operating conditions. A regulatory inquiry may reveal governance issues weakened in a company's own narrative. The effect of a tax preference depends on the subject type and policy details. The sentiment of a financial news headline cannot be equated with a market judgment. For this reason, the course design places how students form conclusions before what conclusions they submit. In project assignments, students are required to present problem definition, sources, prompt design, summaries of model outputs, verification evidence, and revision notes.

2.2. Implications of Financial Application Literature for Course Content

Research on artificial intelligence in finance provides content resources that can be transferred into the course. Studies on financial large language models, financial benchmarks, financial text mining, financial statement analysis, and financial advice show multiple paths through which models enter financial data and financial tasks [12-24]. FinGPT highlights open-source financial large language models and financial data workflows. BloombergGPT shows the possibility of training on both financial and general corpora. FinBen discusses the evaluation of financial large language models from the perspective of benchmark testing [15-18]. These studies help students understand three issues: why models need domain materials, why financial tasks require specialized evaluation, and why model outputs cannot be judged only by fluency.

At the same time, the literature on financial stability and financial governance reminds the course to place risk boundaries in an important position. The European Central Bank has discussed the potential benefits and risks of artificial intelligence for financial stability, and the Financial Stability Board has also drawn attention to model risk, operational risk, and market concentration risk that may arise from artificial intelligence systems [13-14]. Remolina analyzes hallucination, bias, accountability, and transparency in financial generative AI from legal, ethical, and technological governance perspectives [22]. These studies have direct classroom implications. When students process annual reports, they need to check whether the model has omitted important passages or treated standardized disclosure as substantive risk. When comparing policies, they need to verify release dates, applicable subjects, effective scope, and wording of clauses. When reading investment research materials, they need to distinguish instructional analysis from investment advice. When using public data, they also need to discuss which materials may be uploaded to external models, which should be desensitized, and which should not be uploaded.

The course therefore arranges financial application and risk governance along the same teaching line. Models are placed inside the financial analysis process, and their role is constrained by evidence, calculation, and responsibility boundaries. Students learn prompt design, structured extraction, and model comparison, while also learning source verification, data compliance, and professional interpretation. In this arrangement, generative artificial intelligence is not a tool-training topic detached from professional courses. It becomes a constrained method through which finance-related graduate students understand materials, organize evidence, and express conclusions.

3. Course Resource Development Framework

3.1. Course Positioning and Construction Basis

Generative Artificial Intelligence and Financial Applications is positioned as an applied course for graduate students in finance-related disciplines. Its main task is to connect AI-assisted methods with the analysis of financial problems. The course was not built from scratch. Earlier Python-based finance teaching had been carried out for years in graduate training related to asset appraisal, taxation, public finance, and other areas. Project application materials show that the teaching team had already formed earlier reform foundations in large-model-enabled classrooms, blended teaching of financial data and Python, and asset appraisal course construction. The team had also accumulated a textbook on financial big data analysis using Python and MOOC resources in Python and financial data analysis as well as asset appraisal. These foundations allow the course to start from authentic financial materials and students' existing data-learning experiences, reduce the share of purely introductory tool teaching, and devote more classroom time to financial task design and output verification.

After the course was renamed, the earlier teaching foundation had to be reorganized. Many students had already encountered Python, data cleaning, financial indicators, and public finance cases. Yet they still lacked training in using generative models for long-text processing, structured extraction, prompt design, retrieval support, output verification, and professional interpretation. The course objectives were accordingly set in four areas: understanding basic concepts such as large language models, prompts, retrieval-augmented generation, multimodal models, hallucination, and domain models; practicing summarization, structured extraction, model comparison, code assistance, and simple data organization; placing model outputs back into financial contexts and interpreting them with rules, indicators, and professional knowledge; and using source records, verification logs, and revision notes to make learning processes and responsibility boundaries traceable.

3.2. Components of the Resource System

The course resources consist of conceptual lectures, case materials, experimental task packages, reading assignments, process portfolios, verification logs, and assessment rubrics. Conceptual lectures provide basic explanations. Their content includes the basic working logic of large language models, common elements of prompt design, the idea of retrieval-augmented generation, multimodal material processing, model hallucination, and domain models. The lectures do not aim to cover technical details exhaustively. They serve financial tasks, so that students can understand why model outputs need verification and why the same task may produce different results under different source constraints.

Case materials provide professional scenarios. They mainly come from public annual reports, listed-company announcements, regulatory inquiry letters, fiscal and tax policy documents, financial news, industry reports, and public databases. Material selection follows three principles. Sources should be public and easy to verify; problems should be clear and suitable for classroom discussion; difficulty should increase gradually so that training can move from single skills to integrated projects. Experimental task packages turn case materials into operable learning tasks. They usually include objectives, background materials, source descriptions, operating steps, submission requirements, and reflection questions. Reading assignments bring literature into course discussion. When students read literature on educational assessment, financial generative AI, and AI governance, they need to identify research questions, methods, data, findings, limitations, and the ways in which these studies can be translated into course cases.

Process portfolios, verification logs, and assessment rubrics serve the design of assessment. A process portfolio records how students choose sources, design prompts, handle model outputs, and revise conclusions. A verification log asks students to explain which facts have been checked against original texts, which numbers have been recalculated, and which policy conditions affect the judgment. The rubric shifts the teacher's attention from whether the final report reads smoothly to whether the problem is clear, sources are reliable, model assistance is appropriate, financial interpretation is valid, and process records are complete. These resources form the operational layer of the course. They show students how to complete tasks and give teachers a basis for questioning and assessment.

In resource development, the course gives special attention to the fit among materials. Conceptual lectures that are separated from cases easily become general technical introductions. Case materials without task packages may leave students at the level of excerpting and summarizing. Assessment rubrics without process portfolios still force teachers to evaluate only the final text. For this reason, course resources are organized along a chain of conceptual explanation, case entry, experimental operation, process recording, classroom questioning, and assessment feedback. Each experimental task is designed, as far as possible, to correspond to one verifiable material set, one group of model-operation requirements, one verification checklist, and one reflection question. With this arrangement, students do not merely memorize a prompt template. They gradually become familiar with the full path from source selection to conclusion expression in financial materials.

The course also retains a resource renewal mechanism. Generative AI tools, policy documents, regulatory materials, and financial cases all change. Course resources should not remain fixed after one round of construction. Stable resources include course objectives, competency requirements, task types, rubrics, and process-record templates. Updating resources include currently available models, annual cases, latest policy texts, industry reports, and classroom demonstration materials. Before each teaching cycle, teachers can update cases and tool descriptions while maintaining continuity in competency goals and assessment standards. For a graduate course, this distinction between stable and updating resources helps the course accumulate over time and prevents classroom content from being driven by a single platform or a single year's topic.

3.3. Organization of Content Modules

The course content is organized into four modules: basic understanding, method training, financial scenarios, and risk governance. The basic understanding module helps students distinguish generated content from verified facts. It emphasizes that the linguistic coherence of model output does not automatically become factual reliability. The method training module includes prompt design, structured output, retrieval support, model comparison, code assistance, and data organization. Prompt design is not taught as an abstract technique in isolation. It is embedded in tasks such as policy comparison, annual-report risk extraction, and financial statement interpretation. Students are required to explain source boundaries, comparison criteria, output forms, and verification steps.

The financial scenario module is the main body of the course. It covers annual report analysis, listed-company announcements, tax policy comparison, financial news interpretation, risk disclosure, green finance texts, valuation materials, and investment research reports. Public materials are used by default. The risk governance module runs through these scenarios and addresses academic integrity, data compliance, model bias, verification of generated content, financial consumer protection, and responsibility boundaries. The four modules intersect in teaching. A policy comparison task trains prompt design and clause verification at the same time. A financial statement task involves concept explanation, model use, calculation checking, and professional judgment. A task that reviews an investment research report involves sources, analytical framework, responsibility boundary, and expression norms. This cross-structure helps prevent the course from being divided into isolated tool demonstrations.

3.4. Development of Six Types of Case Resources

The case library is built around six types of tasks. Each type corresponds to clear materials and financial problems. The first type involves annual reports and announcements, mainly training long-text reading. Students work on management discussion and analysis, risk disclosure, audit opinions, or regulatory inquiry letters. Models may assist extraction and preliminary classification, but final judgments must return to the original text to check omissions, vague classification, and unsupported expressions. These cases are suitable for showing how model outputs in long-text processing may appear complete while missing decisive conditions.

The second type involves financial statements and uses calculation to test language. Students may ask a model to draft interpretations of profitability, solvency, operating efficiency, or cash flow quality. Those interpretations must then be checked against formulas, units, accounting definitions, and industry context. The course requires students to distinguish the model’s textual explanation, indicators they have recalculated, original numbers in public financial statements, and interpretive differences arising from accounting-caliber changes. This training prevents students from treating general financial comments generated by a model as professional analysis.

The third type involves policy texts. It trains students to handle applicable subjects, key clauses, effective periods, and possible effects in fiscal policy, tax policy, financial regulatory policy, and digital economy policy documents. The difficulty of policy materials lies in their dense conditions. A model may summarize policy objectives while omitting scope of application, exclusion conditions, or time boundaries. Through policy comparison tasks, the course asks students to list clause bases and judgment conditions, so that model summaries are verified item by item.

The fourth type involves financial news and public opinion. It requires students to distinguish textual sentiment from financial meaning. News headlines and comment tones can help students identify market narratives, but they cannot replace financial judgment. Students need to place news in relation to disclosure timing, market expectations, company announcements, and industry background. In these cases, the course stresses that positive or negative labels produced by a model are only preliminary text-level classifications. Subsequent analysis must return to the nature of the event and the evidence chain.

The fifth type involves investment research. Students use public information and model assistance to prepare report frameworks or initial analyses, then revise them through source verification, indicator logic, and professional reasoning. The course clearly limits these reports to instructional exercises and does not present classroom analysis as investment advice. Students need to explain which information comes from public sources, which structures were assisted by a model, which judgments were manually verified, and which conclusions remain limited by available materials.

The sixth type involves risk governance. The course treats errors themselves as teaching materials. Examples include a model misreading a table, citing a policy incorrectly, producing inappropriate financial advice, exposing sensitive information, or treating general disclosure as material risk. By analyzing these errors, students understand why model outputs must be jointly constrained by sources, calculation, context, and responsibility boundaries. The case library also sets a gradient of difficulty. Early tasks use short policy notices and simple extraction tasks. Mid-course tasks arrange annual-report comparison and financial-statement interpretation. Later tasks use more open-ended assignments such as industry report review or policy impact analysis.

4. Teaching Practice Design

4.1. Classroom Organization

Classroom teaching combines explanation, demonstration, practice, and discussion. Explanation provides conceptual grounding. Demonstration shows how models behave when they encounter financial materials. Practice allows students to complete concrete operations. Discussion tests whether model outputs can withstand professional questioning. Demonstrations deliberately include errors. A teacher may ask a model to explain a financial table and then ask students to check the ratio definitions. The teacher may also ask a model to summarize a policy document and then ask students to find missing applicable conditions. An answer that exposes problems is often more useful as classroom material than an answer that looks complete, because it shows where students need to intervene and revise.

Classroom discussion does not rush toward a standard answer. Students first mark problems in model outputs, compare judgments within groups, and then receive teacher comments. This organization fits financial problems because financial conclusions often depend on specific contexts and relationships among materials. If students look only at a model’s conclusion, they may miss the connections among announcements, audit opinions, policy clauses, and data caliber. Through discussion and questioning, model outputs are turned into learning materials that can be verified and explained.

4.2. Teaching Process

A teaching unit usually begins with the introduction of a financial problem, such as identifying corporate risk, judging policy effects, reading market news, or explaining items in financial statements. The problem should be concrete enough to direct students' attention to financial judgments that materials can support, rather than letting them treat the classroom task as asking a model to summarize materials. The next step is task decomposition. Students decide which parts require source retrieval, which parts can use model assistance, which parts require calculation, and which parts depend on professional judgment.

The first teaching step is task design. Students clarify source boundaries, prompt structure, output form, comparison criteria, and verification checklists. Source boundaries constrain which materials the model may use. Prompt structure specifies how the model should process the materials. Output form makes later verification easier. Comparison criteria help students avoid vague comparison. The second step is model-assisted operation. Students keep prompt records and output summaries, and mark facts, numbers, policy clauses, and interpretive judgments that need verification.

The next step involves material verification and presentation questioning. Students explain their conclusions and answer questions about sources, indicators, policy conditions, and the revision process. Questions from the teacher and peers push students back to the materials and calculations. Students need to explain which parts come from model suggestions and which judgments have been manually verified. The final step is reflective revision. Students write brief reflections on what the model helped with, what errors appeared, what they verified by hand, and how the conclusion changed after verification. This process prevents students from treating model output as a conclusion too early. It also clarifies the path among problem, material, model, verification, and interpretation.

The process can be adjusted across cases. Policy text tasks usually spend more time on source definition and clause verification because applicable conditions define the boundaries of conclusions. Financial statement tasks need an additional calculation-checking link, in which students explain indicator formulas, data sources, and accounting caliber. Financial news tasks put more weight on background supplementation and event interpretation, so that students do not remain at the level of sentiment classification. Teachers can adjust the time for each link according to the task type, but the problem introduction, verification, and reflection links should not be omitted. Problem introduction keeps the class centered on financial judgment. Verification subjects model output to evidence. Reflection helps students turn one operational experience into transferable methodological awareness.

This process also changes the way feedback is provided. Traditional assignment feedback often focuses on the structure, language, and conclusion of the final report. By the time students receive the feedback, it is already difficult for them to return to the original materials and revise. Process-oriented teaching moves feedback forward to the stages of prompt design, source selection, and verification checklist construction. Before students form final conclusions, teachers can point out that the material scope is too narrow, the prompt goal is ambiguous, the output format is not easy to verify, or policy conditions have been omitted. Such feedback is closer to research training, since graduate students need to learn not merely how to write a conclusion at the end, but also how to detect insufficient evidence during analysis and adjust the path in time.

4.3. Two-Round Tasks and Reverse Review

Two-round tasks are frequently used in the course. In the first round, students use a model directly and observe the output. In the second round, they repeat the same task under constraints on sources, required output structure, and verification checklists. The two results are discussed together. Students can see how loose prompts produce answers that look reasonable but have weak foundations. They can also see how source control and verification requirements change conclusions. In a policy comparison task, for example, the first-round output may only summarize policy objectives, while the second round requires the applicable subjects, effective time, core clauses, and exclusion conditions to be listed. The difference itself becomes teaching material.

Group projects extend this logic. A semester project may focus on annual-report risk identification, tax policy analysis, financial public-opinion organization, financial statement interpretation, or review of an investment research report. Outputs include a report, a process portfolio, and a presentation. The process portfolio records sources, prompts, output summaries, verification evidence, and revision notes. During the presentation, students explain what the model completed, what judgment the group made, and which evidence supports the conclusion. The course also uses reverse review. After one group submits a project, another group examines its sources, verification notes, and chain of reasoning, then asks questions about source authority, indicator interpretation, policy conditions, and conclusion logic. When students review others' work, they often understand normative requirements more easily and can see evidence gaps in their own projects.

4.4. Differentiated Support and Classroom Pace

Students enter the course with different technical foundations. Some have programming experience and can complete data processing and simple automation tasks. Others mainly use model interfaces for writing, retrieval, or material summarization. The course therefore uses differentiated tasks. All students complete basic tasks such as text summarization, structured extraction, policy comparison, and simple code assistance. Students with stronger foundations can further undertake retrieval-supported tasks, small knowledge bases, model comparison, or automated data processing. Group projects can also divide work among source collection, model interaction, data processing, financial interpretation, and normative review.

Differentiation mainly concerns technical paths, while academic requirements remain consistent. Every student must define problems, verify materials, interpret results, and keep process records. At the beginning of the course, a short diagnostic task is used to understand students' preparation in programming, data processing, and model use. The diagnostic result supports grouping, task arrangement, and resource allocation, and is not included in grading. Classroom pace should also avoid being led by tool operations. Tool demonstrations should serve financial problems. Classroom discussion should return to evidence and interpretation. After-class tasks should remain constrained by process portfolios and verification logs.

5. Key Issues and Safeguard Mechanisms

5.1. Tool Updates and Reliability of Model Outputs

Generative AI tools change rapidly. A course built around the interface of one platform may become outdated quickly. This course uses relatively stable financial tasks as its organizing center, including summarization, extraction, policy comparison, statement interpretation, risk identification, and report review. Tools may be updated, while task logic, verification requirements, and assessment standards need continuity. Course documents are also divided into stable and updating parts. The stable parts include course objectives, task types, rubrics, and normative requirements. The updating parts include model tools, current cases, policy documents, and industry materials. This arrangement is consistent with earlier experience in finance data and Python teaching, since code tools and data sources have always required periodic updating.

The reliability of model output needs repeated training in the classroom. Models may fabricate facts, confuse concepts, omit conditions, or write in a tone that exceeds the degree of evidential support. The course includes these typical deviations in the case library. Students are required to return to the original text to find omissions or misreadings, return to data to check units, definitions, and calculations, and return to financial logic to judge whether a conclusion fits the institutional background and the specific case. The assessment also recognizes the value of finding errors. A report that identifies and corrects model errors with evidence is more consistent with course objectives than a fluent report without verification records. The rubric therefore asks students to point out weaknesses in model outputs or explain a specific instance in which verification changed the conclusion.

5.2. Academic Integrity and Assessment Validity

Generative artificial intelligence has changed the way course assignments are completed. It has also weakened teachers' ability to judge student effort from the final text alone. The course therefore translates academic integrity requirements into process requirements. Francis et al. discuss the tension between innovation and integrity after generative artificial intelligence enters higher education [10]. The Artificial Intelligence Assessment Scale proposed by Perkins et al. also suggests that courses can define different levels of tool-use boundaries according to learning objectives [11]. Students may use models within prescribed limits to assist material organization, summarization, code prompting, and language revision, but they must record the process and explain how model outputs were verified, modified, and absorbed. Assessment does not focus on tracing whether a particular sentence was generated by a model. It focuses on whether students can explain their sources, judgment basis, and revision process.

This arrangement helps maintain assessment validity. The course aims to cultivate financial problem-analysis ability. Producing a smooth report alone is insufficient evidence that students have completed professional training. Process portfolios, verification logs, and presentation questioning jointly form an evidence chain. They help teachers judge whether students have completed material reading, data checking, policy understanding, and professional interpretation. For students, these requirements also change the way they use models. A model changes from a tool for replacing assignment completion into an auxiliary object that must be recorded, verified, and revised.

The course also needs to distinguish permitted use from inappropriate use. Permitted uses include organizing public materials, generating reading outlines, proposing verification checklists, assisting code debugging, and improving language. Inappropriate uses include directly submitting unverified model outputs, fabricating sources, writing model-fabricated policies or data into reports, uploading materials that should not be shared externally, and being unable to explain the formation of conclusions during presentations. At the beginning of the course, the teacher explains these boundaries and repeats the submission requirements for each project task. This reduces formal disputes and encourages students to develop explainable use habits.

5.3. Data Security and the Centrality of Financial Problems

Finance teaching often involves sensitive materials. The course therefore uses public annual reports, public policy documents, public data, and desensitized teaching data by default. Students may not upload non-public enterprise materials, personal information, client information, or confidential documents to external models. Data security is implemented through judgment exercises. The teacher provides several sample materials and asks students to judge which can be used directly, which require desensitization, and which are unsuitable for use with external tools. Compliance requirements thus enter task operation, rather than remaining as general statements in the syllabus.

Each case must return to a financial problem. What has the policy changed? What does the ratio show? What risk does the disclosure reveal? Which evidence supports the conclusion? These questions determine the direction of model use. At the end of a case class, students are asked to write one financial conclusion in concise professional language. The exercise is short, but it helps students bring tool operations back to evidence and professional interpretation. Course assessment follows the same principle. Technical complexity does not represent learning quality; evidence-based financial analysis is the core.

5.4. Differences in Student Foundations and Insufficient Financial Professionalism

The course must also handle teaching difficulties arising from differences in student foundations. Graduate students in finance-related disciplines come from different backgrounds, with substantial differences in programming ability, data-processing experience, and model-use experience. If the technical threshold is set too high, some students will spend most of their energy on operational details and find it difficult to return to financial problems. If technical requirements are lowered too much, students with stronger foundations will not receive enough extended training. Differentiated tasks and group collaboration can ease this tension, but the teacher needs to make the baseline requirements clear: every student must complete problem definition, source verification, financial interpretation, and process recording.

Insufficient financial professionalism is another issue. Students may produce a report with a complete structure but fail to explain the meaning of indicators, policy conditions, or sources of evidence. The rubric includes the validity of financial interpretation to address this problem. Fluency is not the primary standard. The key is whether the conclusion matches the materials, whether indicator interpretation follows the proper caliber, whether policy judgment lists the conditions, and whether risk warnings are supported by evidence. In classroom questioning, teachers should pursue these issues so that students understand that model output has course value only after professional examination.

6. Observable Dimensions of Teaching Effects

This paper focuses on the design of course resources and processes. Future evaluation can be organized around observable evidence in course construction and student work. The resource dimension concerns whether the course has formed a syllabus, case library, experimental task packages, reading list, assessment rubric, and update records. The process dimension concerns whether students can keep source records, verification logs, and revision notes, and whether they can explain these records in discussion. The ability dimension concerns whether students can raise financial questions, organize source materials, verify claims, and write professional analysis. The normative dimension concerns whether students can identify data boundaries, cite sources, and distinguish model assistance from human judgment.

These dimensions can be translated into future empirical evaluation designs. Process portfolios can be used to analyze whether students record sources, prompts, outputs, and revisions. Presentation scoring can be used to observe whether students can explain their basis under questioning. Reflection reports can show students' understanding of model limitations, verification responsibility, and professional judgment. Classroom observation can record how groups find problems and trace evidence. If a more formal evaluation is to be conducted, measurement tools for AI literacy, financial problem-analysis ability, and normative awareness can be designed before and after the course. Such research needs samples and tools to be prepared before teaching begins; post-course impressions cannot serve as a substitute.

The course's existing teaching foundation makes this work feasible. Earlier finance data and Python teaching has accumulated assignments, cases, code tasks, and experience with cross-disciplinary students. The next step is to connect these foundations with formal evaluation designs for generative AI teaching. For course construction, the more urgent task at the current stage is to stabilize observable evidence so that later quantitative evaluation will have materials to rely on.

In implementation, observable dimensions can also correspond to different evidence sources. Such evidence design does not require an immediate conclusion that effects have improved, but it does provide a relatively stable data basis for later research.

Future evaluation should avoid two misunderstandings. First, the frequency of model use should not be equated with learning effect. More frequent model use does not mean stronger financial analysis ability. Without verification and interpretation, frequent use may even hide weak professional judgment. Second, the fluency of the final text should not become the main indicator. Generative artificial intelligence can significantly improve the surface form of language. Teachers need to observe whether students can explain data caliber, policy conditions, and evidence chains. For these reasons, later evaluation should combine product results, process materials, and classroom questioning to form a fuller judgment.

7. Value for Wider Application

The resource structure of Generative Artificial Intelligence and Financial Applications can support other finance-related courses. Taxation courses can use policy texts and tax administration cases. Finance courses can use market information and risk disclosure. Accounting courses can use financial reports and audit texts. Asset appraisal courses can use industry materials and enterprise value cases. Digital economy courses can use platform data, public policies, and industry reports. What can truly be transferred is the working sequence of source description, output recording, fact verification, professional interpretation, and normative reflection. Different disciplines can replace the materials while retaining requirements for evidence chains and process records.

Application in other courses also has boundaries. Different fields have different data rules, model needs, and professional standards. Policy text cases in taxation cannot simply be copied into accounting classrooms. A shared process must be combined with disciplinary materials, and teachers need to raise questions with professional specificity. For graduate education in finance-related fields, generative artificial intelligence has not reduced the importance of reading, calculation, and argumentation. It has made gaps in these abilities easier to see. If course construction uses this visibility well, model assistance can become an opportunity to strengthen graduate training.

Course-group construction can begin with resource sharing. Case libraries, task packages, assessment rubrics, and reading lists can be reused across different courses, while each course should retain its own professional questions. This reduces repeated construction and avoids turning generative AI course resources into a generic tool tutorial. Graduate course reform oriented toward "AI + Finance" must ultimately serve the cultivation of financial professional competence. The model is only one analytical tool used under norms.

At the college level, several existing case-based courses can first be selected as pilots. The process portfolio template, verification log, and presentation questioning method from this course can be embedded in existing teaching links. A taxation course may begin with verification of policy applicability conditions. An accounting course may begin with reading notes to financial statements and audit opinions. A finance course may begin with risk disclosure and interpretation of market information. These courses do not have to use the same model or the same set of cases, but they can share the basic requirements of source, output, verification, interpretation, and revision. After one or two teaching cycles, the completeness of students' process materials and the quality of problem analysis can be compared across courses, gradually forming a shared resource library for finance-related graduate courses.

8. Conclusion

Generative Artificial Intelligence and Financial Applications is built on an existing foundation of finance data teaching and real classroom needs. Finance-related students already encounter generative artificial intelligence in reading, programming, retrieval, and writing. The course needs to bring this use into a financial analysis process under normative constraints. The resource design proposed in this paper organizes content around financial scenarios, builds cases with public materials, writes experiments as task packages, connects literature reading with case construction, and assesses both the process and the final argument. The teaching process begins with financial problems, moves through task decomposition, model assistance, source verification, presentation questioning, and reflective revision, and finally points toward the formation of professional interpretive ability.

The main construction focus of the course is to help students still explain sources, verification bases, calculation caliber, and financial judgments in a model-assisted environment. Through process portfolios, verification logs, reverse review, and assessment rubrics, the course brings model use into a teaching process that can be discussed, traced, and assessed. Tool operations are thereby returned to professional learning goals. Future work should continue to improve the case library, refine assessment rubrics, and conduct empirical evaluation across multiple teaching cycles.

Acknowledgements

This study was supported by the 2025 "Artificial Intelligence+" Graduate Education and Teaching Reform Research Project of Central University of Finance and Economics (ZCJG202502).

References

[1] Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

[2] UNESCO. (2024). *AI competency framework for teachers*. https://www.unesco.org/en/articles/ai-competency-framework-teachers

[3] Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). A systematic review of generative AI for teaching and learning practice. *Education Sciences*, 14(6), 636.

[4] Qian, Y. (2025). Pedagogical applications of generative AI in higher education: A systematic review of the field. *TechTrends*, 69, 1105–1120.

[5] Wu, F., Dang, Y., & Li, M. (2025). A systematic review of responses, attitudes, and utilization behaviors on generative AI for teaching and learning in higher education. *Behavioral Sciences*, 15(4), 467.

[6] Lee, S. S., & Moore, R. L. (2024). Harnessing generative AI for automated feedback in higher education: A systematic review. *Online Learning*, 28(3).

[7] Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410.

[8] Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education*.

[9] Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, 100151.

[10] Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, 81, 14048.

[11] Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, 21(6).

[12] Lee, D. K. C., Guan, C., Yu, Y., & Ding, Q. (2024). A comprehensive review of generative AI in finance. *FinTech*, 3(3), 460–478.

[13] Leitner, G., Singh, J., van der Kraaij, A., & Zsámboki, B. (2024). The rise of artificial intelligence: Benefits and risks for financial stability. European Central Bank. https://www.ecb.europa.eu/press/financial-stability-publications/fsr/special/html/ecb.fsrart202405_02~58c3ce5246.en.html

[14] Financial Stability Board. (2024). *The financial stability implications of artificial intelligence*.

[15] Li, Y., Wang, S., Ding, H., & Chen, H. (2023). Large language models in finance: A survey. In *Proceedings of the Fourth ACM International Conference on AI in Finance* (pp. 374–382).

[16] Yang, H., Liu, X.-Y., & Wang, C. D. (2023). FinGPT: Open-source financial large language models [Preprint]. arXiv.

[17] Wu, S., Irsoy, O., Lu, S., et al. (2023). BloombergGPT: A large language model for finance [Preprint]. arXiv.

[18] Xie, Q., Han, W., Chen, Z., et al. (2024). FinBen: A holistic financial benchmark for large language models [Preprint]. arXiv.

[19] Kim, A. G., Muhn, M., & Nikolaev, V. V. (2024). Financial statement analysis with large language models [Preprint]. arXiv.

[20] Chen, B., Wu, Z., & Zhao, R. (2023). From fiction to fact: The growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, 21(4), 471–496.

[21] Krause, D. (2023). Large language models and generative AI in finance: An analysis of ChatGPT, Bard, and Bing AI [Preprint]. SSRN.

[22] Remolina, N. (2024). Generative AI in finance: Risks and potential solutions. *Law, Ethics and Technology*, (1), 0002.

[23] Mbanyele, W. (2024). Generative AI and ChatGPT in financial markets and corporate policy: A comprehensive review [Preprint]. SSRN.

[24] Lo, A. W., & Ross, J. (2024). Can ChatGPT plan your retirement?: Generative AI and financial advice. *Harvard Data Science Review*, Special Issue 5.