https://doi.org/10.62836/iaet.v5i1.0005

Global Science Publishing

IAET

Innovations in Applied Engineering and Technology
https://ojs.sgsci.org/journals/iaet

Article

An Industry-Oriented Curriculum Framework for Graduate Education in Quantitative Finance: Integrating Large Language Models and Spatio-Temporal Wavelet Networks

Shan Zhao [1,*], Qingjing Zhong [1], Panke Qin [1], Yun Xin [1], Zhanqiang Huo [1], Guosheng Zhang [2], Haiyang Li [3], Ming Xu [1] and Jingwen Guo [1]

[1] *School of Software, Henan Polytechnic University, Jiaozuo 454000, China*

[2] *Beijing Siding Chuanglian Technology Co., Ltd., Beijing 100068, China*

[3] *Zhengzhou Qimi Information Technology Co., Ltd., Zhengzhou 450000, China*

Abstract: The gap between the standard curriculum of quantitative finance and what is needed in practice has been increasing with the speed of financial artificial intelligence. The graduate students are being asked to be able to perform a single workflow involving structured market data, non-structured financial text, cross-asset dependence, non-stationary signal, reproducible backtesting, and deployable computing environments. This paper develops an industry-based curriculum framework which groups these needs on the basis of competency education and project based learning. The structure of the framework includes four interrelated components that include an industry-task and competency map, a three-tiered modular curriculum, a capstone multimodal forecasting project, and a feedback mechanism of revising the curriculum based on evidence. Factor research can be done using Microsoft Qlib; text representation and sentiment analysis can be performed by means of financial language models; discrete wavelet transform, Wavelet Kolmogorov-Arnold Networks, graph convolutional networks and Stockformer can be used to provide progressively more complex modelling operations; Linux-based deployment will allow for reproducibility, monitoring and engineering practice. It makes the difference between the technical metrics of projects, such as Information Coefficient, ICIR, Sharpe ratio, maximum drawdown, and the demonstration of student competence, such as the provenance of data, comparison of models, reliability of the deployment, reasoning about risk, and technical writing. The proposed curriculum design is reusable in terms of domain knowledge, modeling, engineering, and use of generative AI in responsible manner. Future multi-cohort application and outcome testing should prove its educational efficacy.

Keywords: graduate education in quantitative finance; curriculum reform; industry-oriented education; large language models; spatio-temporal modelling; project-based learning; AI engineering

1. Introduction

The professional graduate education is anticipated to be a combination of good theoretical basis and the ability to address real-world issues within a particular industry or profession [1]. Such expectation is especially

Received: 1 July 2026; Accepted: 15 July 2026.

* Corresponding: Shan Zhao (zhaoshan@hpu.edu.cn)

challenging in quantitative finance as professional workflows are quickly transformed with the introduction of new sources of data, modelling methods, computing infrastructure and regulatory needs. A graduate entering this field may need to come up with an investment question, build auditable datasets, assess factors, model temporal and cross-sectional dependence, manage backtesting bias, implement a reproducible pipeline and explain risk to technical and non-technical audiences.

Traditional curricula tend to treat each of these capabilities as a separate topic. Statistical learning can be taught without production-oriented data management; time-series modeling can ignore relationships between assets; natural language processing can be decoupled from price-volume analysis; model evaluation can focus on prediction loss instead of portfolio risk, reproducibility, and deployment robustness. This compartmentalization makes it challenging for students to grasp how each method fits into the full lifecycle of quantitative research. It also leads to a misalignment between finishing courses and demonstrating professional proficiency.

Competency education is a valuable source of information in this issue since it can be initiated on the basis of the capabilities that are visible instead of a given list of topics [2]. Project-based engineering learning also implies that complex competence may occur during the process of the learners applying knowledge to less structured problems at each level of the problem [3]. This process can be supported by the integration of industry and education, where real-life tasks, technical requirements and other feedbacks will be included in curriculum development [4]. Nevertheless, there is no clear indication of how quickly the quantitative-finance approaches are going to be implemented and evaluated within general frameworks of curriculum reform. On the other hand, the research on financial forecasting models tends to pay attention to the predictive ability of the model and does not give much insight into how the process of the workflow of the underlying model could be transformed into an effective graduate program.

This paper bridges that gap by suggesting an industry-focused curriculum model of graduate studies in quantitative finance. The framework applies a multimodal forecasting project as a teaching carrier instead of introducing the integrated architecture as a new trading algorithm. Its contribution is threefold. First, it charts an end-to-end quantitative workflow to five competency areas: data and factor research, multimodal financial interpretation, spatio-temporal modelling, AI engineering, and risk-aware communication. Second, it structures technical information into a gradual curriculum which relates Microsoft Qlib, financial language models, wavelet methods, graph neural networks and attention-based forecasting. Third, it isolates technical performance of projects and educational evaluation and creates a feedback loop on the basis of learner artifacts, evidence of deployment and external review.

2. Curriculum Design and Methods

2.1. Industry Tasks and Competency Mapping

The framework is designed in a backward-design process. It starts with the activities that take place in an end-to-end quantitative research workflow and then defines the knowledge, skills, and professional judgement needed to accomplish such activities. It does not aim at teaching students how to replicate one model, but rather assist them in explaining and justifying decisions made during each step of a research pipeline.

Five areas of competencies are specified. The field of data and factor research is the study of provenance, cleaning, time alignment, feature building, leakage prevention, etc. The financial interpretation of multimodal involves financial text as an additional source of information and students should be able to differentiate between the representation or measurement of sentiment in a text and the prediction of markets. Spatio-temporal modelling encompasses frequency analysis, inter-asset relationships, temporal correlation and comparative assessment with the less complex models. Engineering AI is environment handling, asynchronous operation, logging, observation, and reproducibility. Risk-conscious communication entails portfolio measures, constraints, risk in model, ethics of data, and dissemination of findings in such a way that it does not overstate its facts. The alignment of industry tasks with curriculum modules, representative student tasks, and assessment evidence is summarized in Table 1.

Table 1. Mapping industry tasks to curriculum modules and assessment evidence.

Industry Task	Curriculum Module	Representative Student Task	Assessment Evidence
Build an auditable research dataset	Qlib data and factor research	Construct and document an Alpha158/ Alpha360-based dataset with temporal alignment checks	Data dictionary, provenance record, leakage checklist, reproducible script
Interpret unstructured financial information	Financial language models	Generate time-stamped text representations or sentiment features and justify the aggregation window	Source log, prompt/model record, error analysis, interpretation note
Model non-stationary and cross-asset structure	Wavelet, WaveKAN, GCN, and Stockformer	Compare baseline and advanced models through controlled additions and ablations	Model card, architecture rationale, ablation table, training log
Evaluate a strategy under realistic constraints	Backtesting and risk evaluation	Report IC, ICIR, turnover, transaction-cost assumptions, Sharpe ratio, and drawdown	Backtest configuration, metric definitions, sensitivity analysis
Deliver a reproducible workflow	AI engineering and communication	Deploy the pipeline in an isolated environment and prepare a verified technical report	Environment file, run log, deployment test, oral defence, reviewed report

The end product is not a map of the neural network but a curriculum. The competencies are defined by industry needs, the modules and learning activities are established on the basis of the competencies, and the artifacts of students can be used as an evidence in the assessment process and to revise the curriculum (Figure 1).

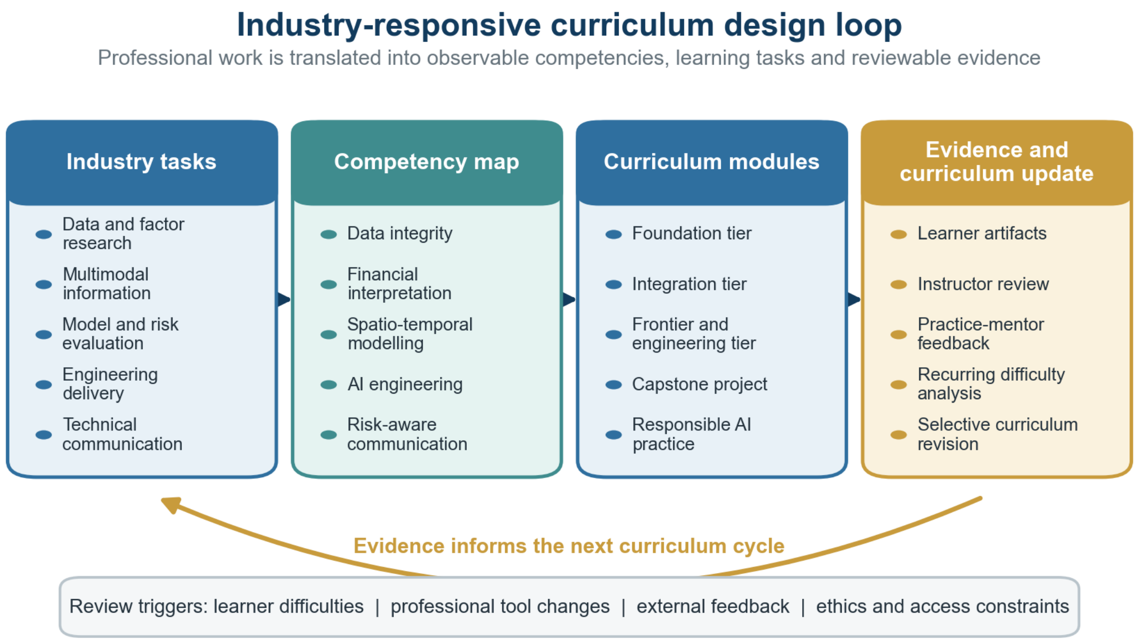

Figure 1. Industry-responsive curriculum design loop for graduate education in quantitative finance.

2.2. Three-Tier Modular Curriculum

The curriculum is structured in three levels that are more technically integrated and learner autonomous. The foundation tier builds financial-statistical reasoning, Python-based data handling, factor construction, baseline modelling, and research reproducibility. Microsoft Qlib is presented at this point since it offers a uniform setting of the preparation of data, factor study, modelling, and backtesting [5]. Students initially replicate transparent baselines like linear models or LightGBM prior to engaging with more complex architectures [6].

The complementary representation of financial information is concentrated on the integration level. The low and high frequency elements of non-stationary series are analyzed using the discrete wavelet transform [7].

Kolmogorov-Arnold Networks, as well as their generalization in terms of wavelets, offer a chance to speak about nonlinear functions that can be learned and sensitive to frequencies [8, 9]. The graph convolutional networks are applied to create relationships between assets where adjacency matrix is based on membership of sectors or supply-chain details or rolling correlations [10]. Stockformer further relates the price-volume information with temporal attention and prediction across sections [11]. Only once the students have been given an initial basis and mentioned the limitation it will attempt to overcome, each component will be incorporated at an advanced stage.

The frontier and engineering level is the introduction of financial language models, multimodal fusion, deployment, monitoring, and responsible AI. Language modelling on a domain-specific scale is given as an example in FinBERT and BloombergGPT [12,13]. The students are to regard the outputs of language-models as variables that need to be validated downstream instead of expecting that there is an automatic correlation between a sentiment score and returns. There are activities such as dependency isolation, remote execution, structured logging, checkpoint recovery and resource monitoring. Log summarization, description of visualizations, or report organization can be done with the help of generative AI, but all numbers, references, and conclusions should be checked by the students.

The Figure 2 is the illustration of the evolution in foundations to modelling and engineering innovation. The vertical axis is an indicator of how much finance, data, modelling, engineering and communication are integrated; the horizontal axis is an indicator of complexity of tasks and independence.

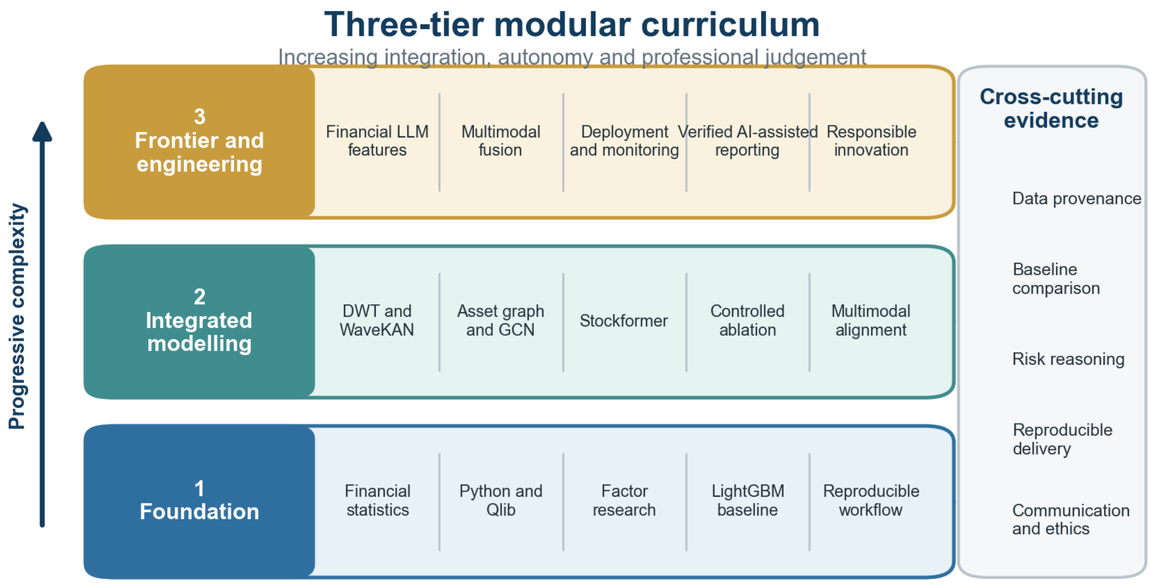

Figure 2. Three-tier modular curriculum for quantitative-finance graduate education.

2.3. Capstone Project Workflow

The capstone project will involve the students in a controlled multimodal forecasting study. A viable project can be based on an equity universe that is documented, a well-defined prediction horizon, price-volume factors, and time-stamped financial text. The initial step is problem formulation: students define the target variable, investment universe, rebalance frequency, and practical constraints. The second level is data governance: students record where each variable came from and when it was available, stop future information getting into training data, and set up train, validation, and test periods.

The third phase is baseline construction. Students apply at least one interpretable or commonly used baseline prior to the inclusion of advanced components. The fourth step is development of progressive models. The wavelet features, graph relations, temporal attention and text features are presented sequentially in order to determine their incremental value. The fifth stage involves backtesting and risk evaluation. Technical measures can be Information Coefficient (IC), Information Coefficient Information Ratio (ICIR), prediction error, turnover, Sharpe ratio and maximum drawdown. These indicators assess the technical artifact; they do not directly measure the quality of learning or teaching.

The last phase is engineering delivery and defence. The students recreate the workflow in a closed system, prove that they can recover after an interrupted run, describe the resource needs and present a short report that relates model behaviour to financial reasoning. When an LLM is employed to write explanations or summarize logs, the submission should have a verification record which explains what was generated, checked, and revised by the student.

2.4. Evidence-Based Assessment and Curriculum Update

Assessment is a combination of process evidence and quality of the final artifact. The data documentation, version history, experimental logs, ablation decisions and responses to review are all examples of process evidence. Technical correctness, reproducibility, risk analysis, deployment reliability and clarity of communication are some of the evidence of artifacts. This difference ensures that a student will not get a high competency judgement simply due to one backtest giving a favourable return.

The course instructor, an enterprise or practice mentor and peer reviewers can be involved in the evaluation process. Instructors evaluate conceptual validity and research design; practice mentors evaluate engineering feasibility and professional standards; peers evaluate reproducibility and communication. The record of assessment is supposed to indicate what evidence supports every judgement. Revision of courses will be informed by recurring issues that have been noted across projects, modifications in industry tools and feedback provided by external reviewers as opposed to novelty of one model. The proposed evidence framework, including assessment dimensions, primary evidence, main reviewers, and assessment points, is summarized in Table 2.

Table 2. Proposed evidence framework for assessing student competence.

Assessment Dimension	Primary Evidence	Main Reviewer	Assessment Point
Data governance and reproducibility	Data dictionary, versioned code, leakage checklist, environment record	Instructor and peer reviewer	Project milestones and final submission
Model design and interpretation	Baseline rationale, architecture explanation, model card, controlled ablation	Instructor and practice mentor	Design review and final defence
Backtesting and risk reasoning	Metric definitions, cost assumptions, sensitivity analysis, limitation statement	Instructor and practice mentor	Final technical review
Engineering delivery	Run logs, recovery demonstration, dependency file, deployment test	Practice mentor and peer reviewer	Demonstration session
Communication and responsible AI	Oral defence, verified report, AI-use and verification statement	Instructor, practice mentor, and peers	Final assessment

It is only in cases of the protocol being approved or exempted that the educational data can be gathered when the results of students are applied to research. The potential indicators that may be used in the future are rubric reliability, project completion and rates of reproduction, successful deployment, pre-course and post-course competency assessment, as well as feedback on the outcome of the course by graduates or employers. These education outcomes should be presented in a different way than technical metrics.

3. Framework Application

3.1. Project-Based Integration of Quantitative Methods

The suggested model translates the set of technical subjects into a single series of choices. The data and experimentation layer is provided by Qlib. A simple model is established as an open reference. Wavelet analysis will take care of non-stationary structure in signal, graph modelling takes care of relationship between assets, temporal attention takes care of sequential dependency, and language-model features that are not included in the price-volume variables. This is due to the fact that the modules are linked to one another via a shared project so the students are required to explain not only the functioning of a method but also its inclusion in the pipeline and how it can be tested.

It is also the differentiated learning that can be facilitated by this sequencing. Quantitative finance novices may have a defensible baseline workflow at the bottom level. More skilled students will use graph building, WaveKAN elements, multi-modal integration or even bespoke risk sensitive goals. The general evaluation model allows to vary the technical approaches but does not change the standards of the data quality, comparison, reproducibility and communication.

3.2. Separation of Technical and Educational Evidence

The three levels of evidence are separated in the core of the framework. The initial one is based on model performance and encompasses IC, ICIR, error statistics, turnover, Sharpe ratio, and drawdown as a complete backtest. The second level relates to the engineering quality and comprises the reproducibility, runtime stability, dependency control, logging, and deployment testing. Thirdly, the student competence involves problem formulation, justification of methodological approach, risk thinking, cooperation, and communication.

These are not the same levels. The good model is strong but can be hidden in data leakage or lack of understanding, and a weak one may have excellent experimental control and critical thinking. It also assesses the quality of research process as well as the numerical results that are obtained as part of the curriculum. Figure 3 provides an overview of the capstone workflow and what was produced at every step.

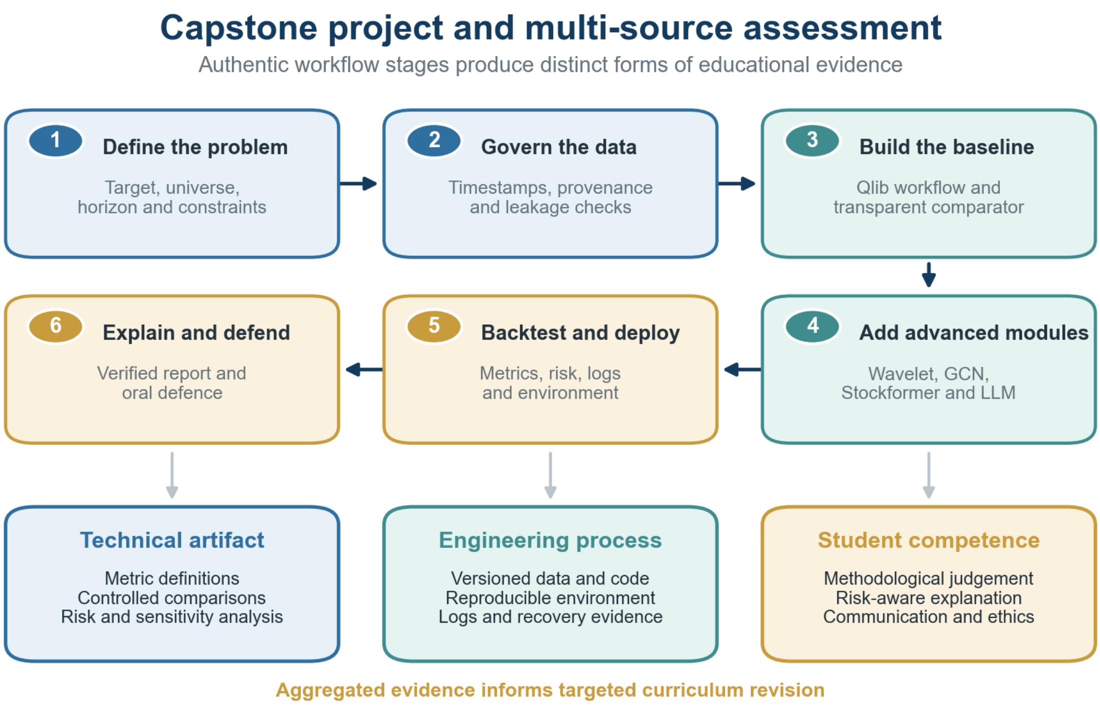

Figure 3. Project-based learning and multi-source assessment workflow.

3.3. AI Engineering and Responsible Reporting

The execution on the basis of Linux, Python environments in isolation, training that is not synchronous and remote development will be preserved as practical material since it introduces students to the situations where quantitative models are created and supported. Such works must not be regarded as research work but as engineering skills. The test should be based on how another person would have been able to recreate the environment, read through the logs, restart a failed one, and check the outputs provided.

The AI-assisted reporting is also considered as a guided learning support. A report outline can be arranged by an LLM, log messages translated into readable descriptions, or questions to analyze errors can be proposed. It must not make performance claims that are not verified, citations, and academic paper that is ready to be submitted on the behalf of the student. The need to provide an AI-use and verification statement makes the process of reporting visible and transforms generative AI into an object of professional judgement instead of an opaque shortcut.

4. Discussion

The suggested program of study is the solution to one of the common issues in the technical graduate studies, where the higher level of techniques is introduced as a separate topic and the learning process is not reorganized or evaluated. This weakness in quantitative finance has been compounded by the interplay of financial theory, non-stationary data, machine learning, computing infrastructure and risk management. The framework replies with an end-to-end project that integrates all these things and the evidence of competence is made visible through objects which can be checked and repeated.

The educational use of the LLMs and spatio-temporal wavelet models can be useful when it is able to produce some interesting contrasts. The students will have an opportunity to compare between structured and unstructured information, local and cross-asset dependence, low- and high-frequency information, as well as simple and complex information. The learning does not rely on all the sophisticated elements that are working to enhance the performance. Even in cases where a component fails under the conditions of controlled testing, the students may still learn about the failure, they would then determine the limits of the failure, and report its results properly.

It also restricts the statements of effectiveness. The paper does not provide a controlled and student-level outcome assessment, it is merely a description of the logic of the curriculum design and implementation. It thus cannot be proven that the framework will enhance the employment, capacity to innovate or performance in investments. This needs to be based on evidence of cohorts of students in the actual course, with open tools and ethical procedures. In the future, the curriculum should be tested among various groups of people, the competency rubric reliability, the comparison between other teaching orders, and the possibility of applying the acquired skills in internships, research work, or job positions.

There are a number of practical risks that should be addressed. Financial data sets can have licensing constraints, survivorship bias and time-alignment errors. LLM services bring in cost, privacy, model-version and hallucination risks. The complex architectures may promote the shallow stacking of models and can put students with limited access to computing at a disadvantage. These risks can be reduced by using shared computing resources, baseline-first assessment, explicit data-governance requirements, cost reporting, and lower-compute project pathways.

Lastly, it is important to be selective in the updating of curriculum. AI is not a new field and there are no reasons to abandon old statistics or finance or research design with a short-term solution. Only when a new content can help develop a certain competency, which can be taught by using the resources available, and which can be measured by evidence, then it should be adopted. This is also the way that this principle allows the curriculum to stay active without being unstable.

5. Conclusions

The paper is an industry-focused curriculum design of graduate education in quantitative finance. It relates the work activities to five competence areas, it structures Qlib, financial language models, wavelet analysis, graph modelling, temporal attention and AI engineering into a sequence of learning activities and the capstone project produces evidence that can be assessed. The framework does not consider positive backtest results as an indication of educational success because it has divided the technical model measurements and engineering quality and student skillfulness. The concept will have a practical application in the implementation of the curriculum but the impact on the education should be determined by multi-cohort assessment which is ethically regulated.

There was no new dataset of students that was created and analysed in this curriculum-framework research. The technical data sets that will be used in the subsequent course implementations ought to be reported on the source, version, access terms, preprocessing process and time frame.

Funding

The study was funded by the Henan Province Graduate Education Reform and Quality Enhancement Project, "Machine Learning" (Approval No. YJS2025XQLH10), the Henan Provincial Key Science and Technology Research Program (No. 262102211082), and the Postgraduate Education Reform and Quality Improvement Project of Henan Province (No. YJS2026ZYJC12).

Author Contributions

Writing—original draft, S.Z., Q.Z., P.Q., Y.X., Z.H., G.Z., H.L., M.X. and J.G.; writing—review and editing, S. Z., Q.Z., P.Q., Y.X., Z.H., G.Z., H.L., M.X. and J.G. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

There was no new dataset of students that was created and analysed in this curriculum-framework research. The technical data sets that will be used in the subsequent course implementations ought to be reported on the source, version, access terms, preprocessing process and time frame.

Conflicts of Interest

The authors declare no conflict of interest.

References

1 He Z. The Difference between Academic Degree and Professional Degree. *Chinese Graduate Student* 2007; **2**: 52.

2 Wyne MF, Zhang L, Farahani A. A Review of Competency Based Education (CBE). In Proceedings of the 2021 International e-Engineering Education Services Conference, Petra, Jordan, 22–23 June 2021; pp. 127–132.

3 Lavi R, Bathe M, Hosoi A, *et al.* The NEET Ways of Thinking: Implementing Them at MIT and Assessing Their Efficacy. *Advances in Engineering Education* 2021.

4 Liu Y, Xiang G, Wang J. Research on the Integration Model of Industry and Education in Applied Undergraduate Colleges and Its Influencing Factors. *China Higher Education Research* 2015; **5**: 64–68.

5 Yang X, Liu W, Zhou D, *et al.* Qlib: An AI-Oriented Quantitative Investment Platform. *arXiv* 2020. arXiv: 2009.11189.

6 Ke G, Meng Q, Finley T, *et al.* LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems 2017*; **30**.

7 Mallat S. *A Wavelet Tour of Signal Processing*, 2nd ed.; Academic Press: Cambridge, MA, USA, 1999.

8 Liu Z, Wang Y, Vaidya S, *et al.* KAN: Kolmogorov-Arnold Networks. *International Conference on Learning Representations* 2025; **2025**: 70367–70413.

9 Bozorgasl Z, Chen H. Wav-KAN: Wavelet Kolmogorov-Arnold Networks. *arXiv* 2024. arXiv:2405.12832.

10 Kipf TN, Welling M. Semi-Supervised Classification with Graph Convolutional Networks. *arXiv* 2016. arXiv:1609.02907.

11 Ma B, Xue Y, Lu Y, *et al.* Stockformer: A Price-Volume Factor Stock Selection Model Based on Wavelet Transform and Multi-Task Self-Attention Networks. *Expert Systems with Applications* 2025; **273**: 126803.

https://doi.org/10.1016/j.eswa.2025.126803.

12 Araci D. FinBERT: Financial Sentiment Analysis with Pre-Trained Language Models. *arXiv* 2019. arXiv: 1908.10063.

13 Wu S, Irsoy O, Lu S, *et al.* BloombergGPT: A Large Language Model for Finance. *arXiv* 2023. arXiv:2303. 17564.