Jun ZHOU, Chaochao CHEN, Longfei LI, Zhiqiang ZHANG, Xiaolin ZHENG, 2022. FinBrain 2.0: when finance meets trustworthy AI. *Frontiers of Information Technology & Electronic Engineering*, 23(12):1747-1764.
https://doi.org/10.1631/FITEE.2200039

FinBrain 2.0: when finance meets trustworthy AI

Key words: Artificial intelligence in finance; Trustworthy artificial intelligence; Risk management; Fraud detection; Wealth management

Corresponding author: Xiaolin ZHENG
E-mail: xlzheng@zju.edu.cn
ORCID: https://orcid.org/ 0000-0001-5483-0366

Motivation

1. Artificial intelligence (AI) has become an indispensable part of everyday life. However, many people are concerned about the trustworthiness of current AI. To apply AI in business, fundamental issues such as black-box algorithms, decision bias, and nonrobust models have become key factors affecting the further development of Fintech.

2. The need for developing trustworthy AI is urgent due to the potential risk of applying AI in key decision areas.

3. The main objective of this paper is to explore the potential and promise of building a trustworthy financial brain in the upcoming era of trustworthy AI.

Main idea

1. In addition to commonly desired attributes, such as model accuracy, financial services demand trustworthy AI with properties that have not been adequately realized, including interpretability, fairness and inclusiveness, robustness and security, and privacy protection.

2. We review the recent progress and limitations of applying AI to various areas of financial services, and then introduce FinBrain 2.0, a research framework toward trustworthy AI. We also give a discussion on several open issues to provide a comprehensive understanding of the challenges and future directions.

Method

1. We first provide a quick introduction and analysis to common intelligent approaches to financial systems.

2. Then, we briefly review recent advances in AI-powered financial services, including risk management, fraud detection, wealth management, personalized services, and regulatory technology (RegTech).

3. Finally, we combine trustworthy AI with financial AI, presenting a financial research framework called FinBrain 2.0.

Major results

Currently, many practical financial applications have nonlinear, complex, and uncertain behaviors that change dynamically over time. Similar to DL explosive adoption in other fields, DL-based approaches have been widely used in financial scenarios.

Table 1 Summary of the advantages and limitations of typical AI methods

Method	Advantage	Limitation
Rule-based expert systems	Good interpretability, great readability, and fast model inference	Usually not the best performers in terms of prediction quality
Genetic algorithms (GAs) and evolutionary algorithms	Good choice for a large-scale/wide variety of optimization problems, and can find good-quality solutions in a short time of computation	GAs might not find the most optimal solution to the defined problem
Traditional machine learning algorithms	Easy to explain, not bad execution speed, and good performance	Feature engineering required and prone to overfitting
Deep learning based approaches	Less feature engineering required, good handling of multi-dimensional/multivariety data, and great model performance	High time/large amount of data required to train, hard to tune, and difficult to interpret

Major results (Cont'd)

The progress in key areas of financial AI demonstrates the latest breakthroughs in improving service efficiency and reducing costs. However, DL techniques still have shortcomings in terms of privacy, robustness, interpretability, etc.

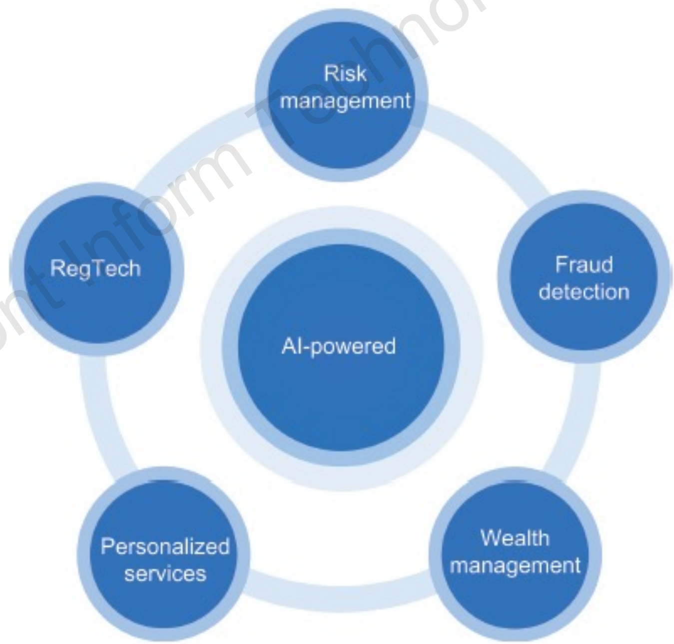

Fig. 1 Artificial intelligence (AI) in financial services

Major results (Cont’d)

The proposed key dimensions of trustworthy AI

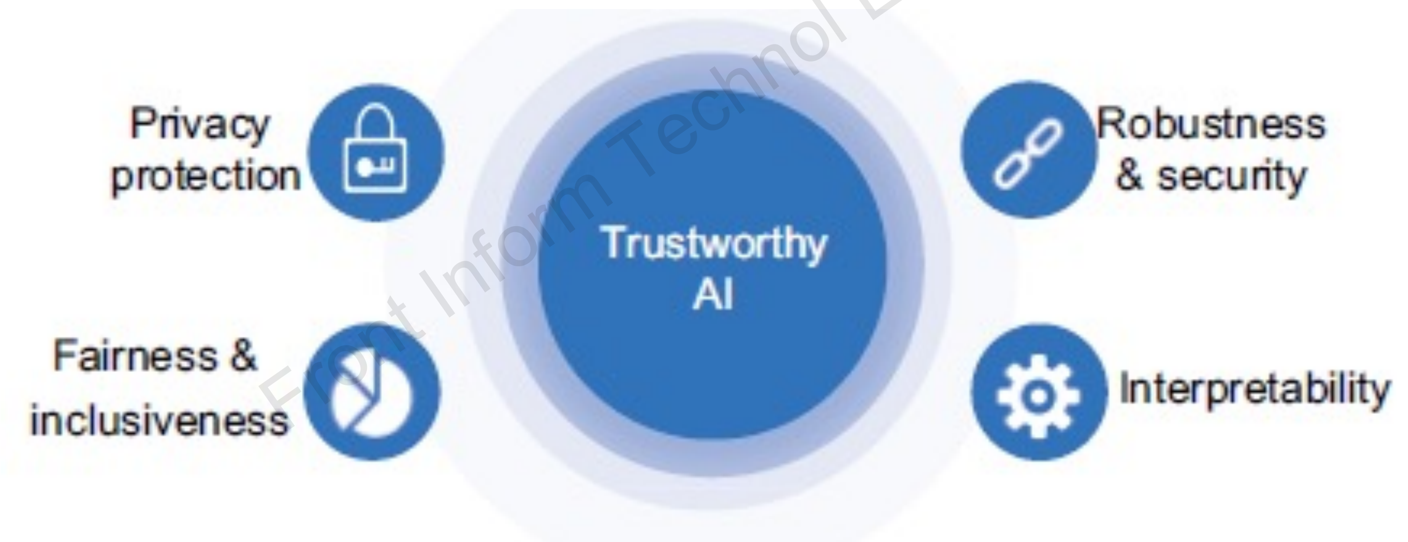

Fig. 2 Four key dimensions of trustworthy artificial intelligence (AI)

Major results (Cont'd)

A thorough literature review on key areas of financial AI

Table S2 Representative financial AI practices

Financial service	Specific item	Data type	Method	P	R	I	Literature
Risk management	Credit scoring	Semistructured	LR	W	W	S	Djeundje et al. (2021)
		Unstructured	RNN, GRU, GNN	W	W	W	Babaev et al. (2019); Cheng D et al. (2020a); Lee et al. (2021)
	Financial distress prediction	Unstructured	LSTM, CNN	W	W	W	Matin et al. (2019); Gregova et al. (2020); Zhang R et al. (2022)
		Semistructured	Tree-based	W	W	S	Huang and Yen (2019); Du et al. (2020); Mousavi and Lin (2020); Shen et al. (2020); Qian et al. (2022)
	Bankruptcy prediction	Unstructured	CNN, GNN, attention	W	W	M	Hosaka (2019); Mai et al. (2019); Zheng Y et al. (2021)
		Structured	Tree-based	W	W	S	Le et al. (2019); Son et al. (2019); Moscatelli et al. (2020); Perboli and Arabnezhad (2021)
Fraud detection	Credit card fraud detection	Semistructured	GAN, CNN, GNN	W	M	W	Fiore et al. (2019); Hu B et al. (2019); Cheng D et al. (2020c); Zhu et al. (2020); Forough and Momtazi (2021); Zhang X et al. (2021); Zheng W et al. (2021)
		Structured	LR, tree-based	W	W	S	Baesens et al. (2021)
	E-commerce transaction fraud detection	Unstructured	GNN, RNN	W	W	W	Liu Z et al. (2018); Cao et al. (2019); Li Z et al. (2021); Lin W et al. (2021); Wang L et al. (2021)
	Loan fraud detection	Unstructured	GNN, attention	W	W	M	Wang D et al. (2019); Zhong et al. (2020); Xu B et al. (2021)
	Insurance fraud detection	Unstructured	GNN	W	W	W	Chen C et al. (2019a); Liang et al. (2019); Cui et al. (2020)
Wealth management	Portfolio management	Semistructured	RL, LSTM, BERT	W	W	W	Hu Z et al. (2018); Xu Y and Cohen (2018); Soleymani and Paquet (2020); Leow et al. (2021); Rezaei et al. (2021); Shi et al. (2021); Xu K et al. (2021)
	Algorithmic trading	Unstructured	RL	W	W	W	Jia et al. (2019); Lei et al. (2020); Théate and Ernst (2021); Yin et al. (2022)
Personalized services	Recommendation	Unstructured	Graph, attention	W	W	M	Sun et al. (2019); Babaei and Bamdad (2020); Bi et al. (2020b); Cheng L et al. (2020); Huan et al. (2021); Li Y et al. (2021); Zhang K et al. (2021)
	Marketing	Unstructured	GNN	W	W	W	Liu Z et al. (2019); Zhuang et al. (2020); Liu Z et al. (2021); Yu et al. (2021); Chou et al. (2022)
	Customer services	Unstructured	RL	W	W	W	Day et al. (2018); Chen C et al. (2019b); Liu J et al. (2020); Wang Z et al. (2021)
		Unstructured	NN, attention	W	W	M	Xu K et al. (2020); Yang et al. (2021)

Notes:

BERT: bidirectional encoder representations from transformers

CNN: convolutional neural network; GAN: generative adversarial network; GNN: graph neural network; GRU: gated recurrent unit; LR: logistic regression; LSTM: long short-term memory

NN: neural network; RL: reinforcement learning; RNN: recurrent neural network

P: privacy; R: robustness; I: interpretability

S: strong; M: medium; W: weak

Major results (Cont'd)

Possible directions for AI trustworthiness

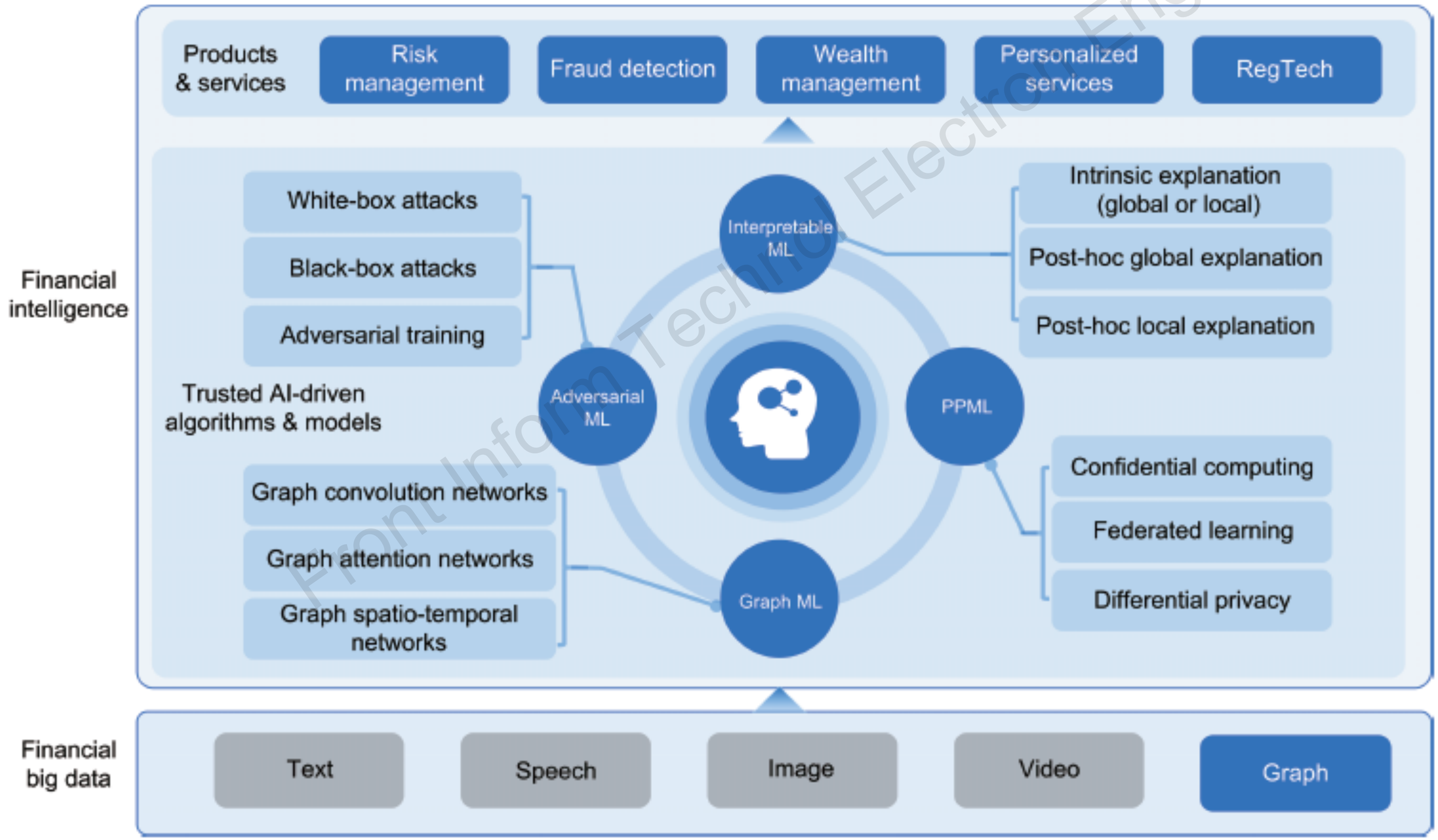

Fig. 3 Overview of a few typical algorithms that contribute to AI trustworthiness

Major results (Cont'd)

The proposed FinBrain 2.0 research framework

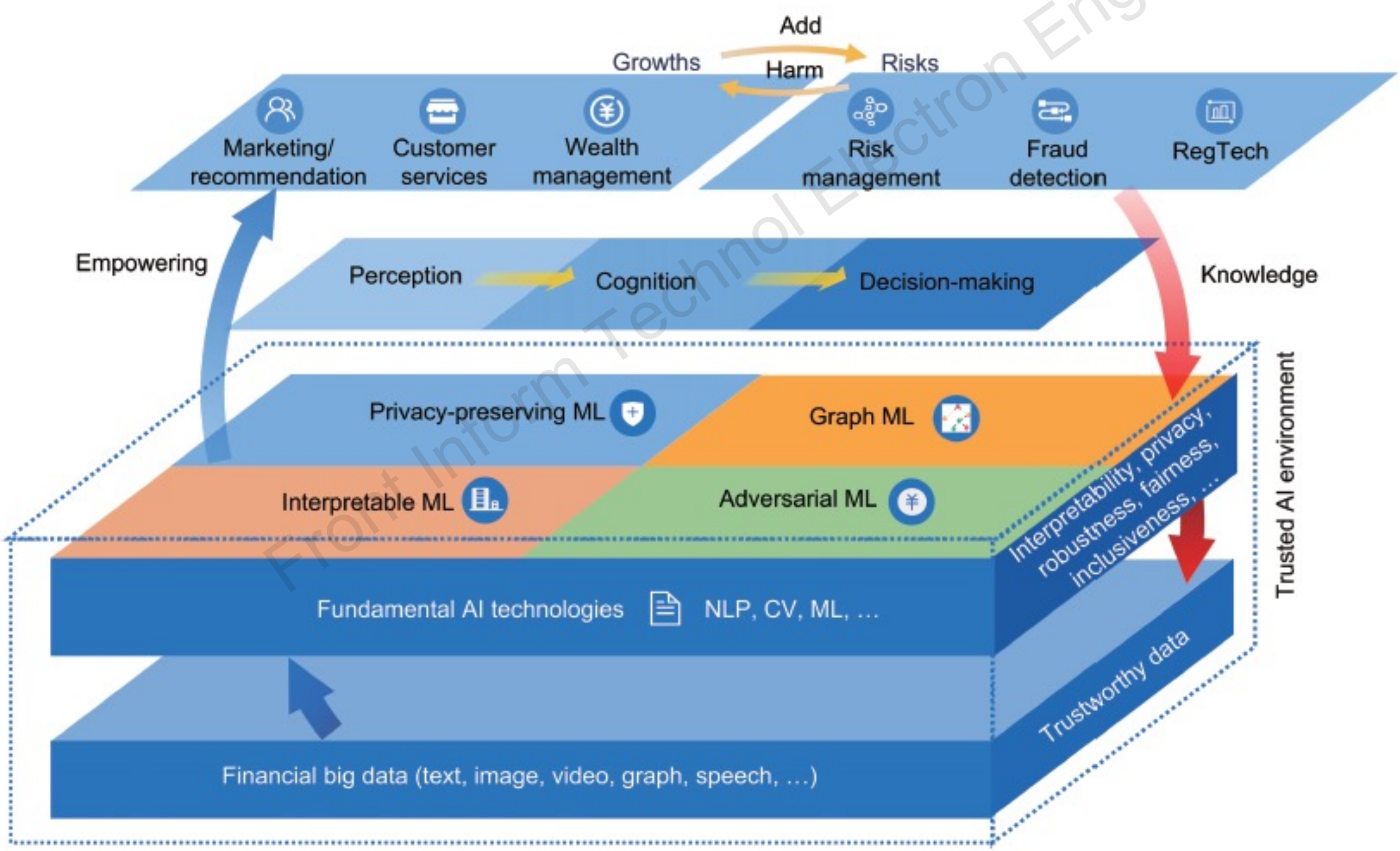

Fig. 4 FinBrain 2.0 framework: the overall research structure

Major results (Cont’d)

Open issues:

1. Financial AI models need to be more flexible and use cross-modal information.

2. Large-scale financial data/model parameters pose a huge challenge to AI systems.

3. Building a trustworthy AI system in finance requires long-term commitment.

Conclusions

1. We review current efforts in the direction of financial AI, including risk management, fraud detection, wealth management, personalized services, and RegTech, and we are only in the early days of this new era, where more research will be implemented and new models will continue to emerge.

2. To overcome the possible flaws in existing financial life, we advocate a shift in focus from performance-driven AI to trust-driven AI, propose a research framework, FinBrain 2.0, and summarize three open issues that attempt to define the future path for prospective researchers.

Xiaolin ZHENG, corresponding author of this paper, is now a professor in the College of Computer Science and Technology, Zhejiang University. His research interests include deep learning, privacy preserving machine learning, and financial AI. He has published more than 70 papers in peer reviewed journals and conferences.

Jun ZHOU is currently a Principal Engineer at Ant Group. His research mainly focuses on machine learning and data mining. He has participated in the development of several distributed systems and machine learning platforms in Alibaba and Ant Financial, such as Apsaras (Distributed Operating System) and KunPeng (Parameter Server). He has published more than 40 papers in top-tier machine learning and data mining conferences.

Chaochao CHEN obtained his PhD degree in computer science from Zhejiang University, China, in 2016, and he was a visiting scholar in University of Illinois at Urbana-Champaign, during 2014–2015. He is currently a Distinguished Research Professor at Zhejiang University. His research mainly focuses on recommender system, privacy preserving machine learning, transfer learning, graph representation, and distributed machine learning. He has published more than 40 papers in peer reviewed journals and conferences.