Digital Finance 3: Technology in Finance

Lesson 25: Introduction to AI/ML in Finance

FHGR

December 7, 2025

Learning Objectives

By the end of this lesson, you will be able to:

- Define artificial intelligence, machine learning, and deep learning
- Understand the hierarchy and relationships between AI concepts
- Identify key applications of AI/ML in finance
- Distinguish between realistic capabilities and overhype
- Recognize the evolution of AI in financial services

What is Artificial Intelligence?

Broad Definition:

- Simulation of human intelligence by machines
- Systems that can reason, learn, and act autonomously
- Originated in 1956 at Dartmouth Conference
- Multiple "AI winters" and resurgences

Key Characteristics:

- Perception (vision, speech)
- Reasoning (logic, planning)
- Learning (from data, experience)
- Natural language processing
- Problem-solving

The AI Hierarchy: From Broad to Narrow

Three Nested Concepts:

1. **Artificial Intelligence** (broadest)
   Any technique enabling computers to mimic human intelligence
2. **Machine Learning** (subset)
   Systems that learn from data without explicit programming
3. **Deep Learning** (subset of ML)
   Neural networks with multiple layers

Analogy:

- AI = Transportation
- ML = Automobiles
- Deep Learning = Electric cars

Modern Reality:
Most "AI in finance" today is actually machine learning, specifically supervised learning algorithms.

Machine Learning: The Core Idea

Traditional Programming:

- Humans write explicit rules
- Input + Rules = Output
- Example: “IF credit score ¡ 600 THEN reject”
- Hard to scale for complex patterns

Machine Learning:

- Algorithm learns rules from data
- Input + Output = Rules (learned)
- Example: Discover credit patterns from 1M loan histories
- Scales to high-dimensional problems

Key Insight:
ML excels when:

- Patterns are complex and non-obvious
- Large amounts of data are available
- Rules are difficult to articulate explicitly

Three Types of Machine Learning

Supervised Learning

- Labeled data (X, Y)
- Learn mapping: X $\rightarrow$ Y
- Examples:
  - Credit scoring
  - Fraud detection
  - Stock prediction
- Most common in finance

Unsupervised Learning

- Unlabeled data (X only)
- Find hidden structure
- Examples:
  - Customer segmentation
  - Anomaly detection
  - Portfolio clustering
- Discovery-oriented

Reinforcement Learning

- Agent learns via trial/error
- Maximize cumulative reward
- Examples:
  - Algorithmic trading
  - Dynamic hedging
  - Game playing (chess, Go)
- Still research-heavy

Deep Learning: Neural Networks at Scale

What Makes It "Deep"?

- Multiple hidden layers (10s to 100s)
- Automatic feature learning
- Inspired by brain neurons (loosely)
- Requires massive data and compute

Breakthroughs (2012-present):

- Image recognition (ImageNet 2012)
- Speech recognition (Google, Apple)
- Language models (GPT, BERT)
- Game mastery (AlphaGo 2016)

Finance Applications:

- Document processing (OCR, contracts)
- Sentiment analysis (news, social media)
- Time series forecasting (limited success)
- Alternative data (satellite, text)

Reality Check:

Deep learning excels with:

- Unstructured data (text, images)
- Millions of training examples
- Pattern recognition tasks

Not always superior for structured financial data (tabular).

AI/ML Applications in Finance: Overview

Risk Management:

- Credit scoring and underwriting
- Fraud detection
- Anti-money laundering (AML)
- Market risk modeling
- Stress testing

Trading and Investment:

- Algorithmic trading strategies
- Portfolio optimization
- Market prediction (limited)
- Robo-advisors
- Alternative data analysis

Customer Service:

- Chatbots and virtual assistants
- Personalized recommendations
- Customer segmentation
- Churn prediction

Operations:

- Document processing (OCR, NLP)
- Regulatory compliance automation
- Process optimization
- Cybersecurity threat detection

Common Thread: Automation of pattern recognition tasks previously requiring human expertise.

Case Study: Credit Scoring Evolution

Traditional Approach (1960s-2000s):

- FICO score (5 factors, fixed weights)
- Linear scorecards
- Based on credit bureau data only
- Transparent, regulated
- Limited predictive power

Limitations:

- Misses non-linear relationships
- Cannot handle alternative data
- One-size-fits-all model

ML Approach (2010s-present):

- Gradient boosting (XGBoost, LightGBM)
- 100s to 1000s of features
- Alternative data (mobile, social, payments)
- Dynamic model updates
- Higher accuracy (10-30% improvement)

New Challenges:

- Explainability ("black box")
- Fairness and bias
- Regulatory acceptance
- Data privacy

Key Lesson: Technology enables better predictions but introduces new risks and ethical questions.

The Hype Cycle: Expectations vs Reality

Gartner Hype Cycle Phases:

1. Innovation Trigger
2. Peak of Inflated Expectations
3. Trough of Disillusionment
4. Slope of Enlightenment
5. Plateau of Productivity

Where is AI/ML in Finance?

- Overall: Slope of Enlightenment
- Deep Learning: Still some hype
- Traditional ML: Plateau (established)
- Generative AI: Peak (2023-2024)

Common Misconceptions:

- "AI will replace all analysts" (No)
- "ML always outperforms rules" (No)
- "More data always helps" (Diminishing returns)
- "Black boxes are always better" (Transparency matters)

Realistic Expectations:

- AI augments, not replaces, humans
- ML excels at narrow, repetitive tasks
- Domain expertise still critical
- Hybrid approaches often best

What AI/ML Can and Cannot Do in Finance

Can Do Well:

- Pattern recognition (fraud, anomalies)
- Classification (credit risk, default)
- Prediction with stable patterns (short-term)
- Data processing at scale (NLP, OCR)
- Optimization (portfolio, pricing)
- Personalization (recommendations)

Success Factors:

- Large, high-quality datasets
- Stable underlying patterns
- Clear objective function
- Ability to validate and test

Cannot Do (or Struggles):

- Predict regime changes (crashes, crises)
- Explain "why" without human input
- Handle novel situations (out-of-sample)
- Replace human judgment entirely
- Guarantee fairness or ethics

Fundamental Limits:

- No free lunch (NFL theorem)
- Efficient Market Hypothesis constraints
- Overfitting to historical noise
- Adversarial dynamics (arms race)

Bottom Line: AI/ML is a powerful tool, not magic. Success requires proper problem framing, quality data, and realistic expectations.

Historical Timeline: AI in Finance

1980s-1990s: Expert Systems

- Rule-based systems (if-then)
- Limited success, brittle
- Example: MYCIN for medical diagnosis

2000s: First ML Wave

- Support Vector Machines (SVM)
- Random Forests
- Credit scoring improvements

2010s: Deep Learning Era

- Neural networks for NLP
- Algorithmic trading explosion
- Robo-advisors launched

2015-2020: Maturation

- Gradient boosting dominance (XGBoost)
- Regulatory frameworks emerge
- Focus on explainability

2020-present: Generative AI

- Large Language Models (GPT-3/4)
- Financial document analysis
- Code generation for analysts
- New regulatory challenges

Future Trends:

- Federated learning (privacy)
- Causal inference integration
- Hybrid human-AI systems

Industry Adoption: Survey Data

Adoption Rates (2023 surveys):

- Large banks: 80-90% have AI initiatives
- Asset managers: 60-70% use ML
- Fintechs: 90%+ (core to business)
- Regional banks: 30-50% (growing)

Top Use Cases:

1. Fraud detection (85%)
2. Customer service chatbots (70%)
3. Credit risk modeling (65%)
4. AML/KYC automation (60%)
5. Algorithmic trading (50%)

Barriers to Adoption:

- Data quality/availability (65%)
- Lack of skilled talent (60%)
- Regulatory uncertainty (55%)
- Integration with legacy systems (50%)
- Explainability requirements (45%)

Investment Trends:

- Global AI in finance market: $10B (2023)
- Projected: $35B by 2030 (CAGR 20%)
- Focus shifting from experimentation to production scaling

Key Players and Ecosystem

Large Tech Companies:

- Google Cloud (AI Platform, AutoML)
- Amazon Web Services (SageMaker)
- Microsoft Azure (ML Studio)
- IBM (Watson Financial Services)

Specialized Fintechs:

- Upstart (AI lending)
- Kasisto (chatbots)
- Kensho (analytics, acquired by S&P)
- Ayasdi (AML)

Traditional Finance + AI:

- JPMorgan Chase (COiN, IndexGPT)
- Goldman Sachs (Marcus, Marquee)
- BlackRock (Aladdin platform)
- Capital One (credit models)

Open Source Community:

- scikit-learn (ML library)
- TensorFlow, PyTorch (deep learning)
- Hugging Face (NLP models)
- Kaggle (competitions, datasets)

Trend: Increasing collaboration between big tech, fintechs, and traditional banks.

Data: The Fuel for AI/ML

Why Data Matters:

- ML models are only as good as training data
- More data often beats better algorithms
- Quality ¿ Quantity (garbage in, garbage out)

Types of Financial Data:

- Structured: Prices, returns, accounting
- Unstructured: News, reports, social media
- Alternative: Satellite, mobile, web scraping
- Real-time: Tick data, order books

Data Challenges:

- Availability (proprietary, expensive)
- Quality (errors, missing values)
- Bias (survivorship, selection)
- Privacy (GDPR, regulations)
- Stationarity (patterns change over time)

Best Practices:

- Rigorous data cleaning
- Train/validation/test splits
- Cross-validation
- Out-of-sample testing
- Monitor data drift

Next Lesson: Deep dive into financial data types and preparation.

Ethical Considerations

Fairness and Bias:

- Models can perpetuate historical discrimination
- Protected attributes (race, gender, age)
- Proxy variables (zip code = race)
- Disparate impact vs. disparate treatment

Transparency:

- Right to explanation (GDPR Article 22)
- Black box models vs. interpretability
- Trade-off: accuracy vs. explainability

Accountability:

- Who is responsible for AI decisions?
- Human-in-the-loop vs. full automation
- Audit trails and governance

Privacy:

- Data minimization principle
- Consent and purpose limitation
- Anonymization challenges
- Model inversion attacks

Regulatory Response: EU AI Act, algorithmic accountability laws, model risk management frameworks.

Skills for AI/ML in Finance

Technical Skills:

- Programming (Python, R)
- Statistics and probability
- Linear algebra and calculus
- ML algorithms and frameworks
- Data manipulation (SQL, pandas)
- Cloud platforms (AWS, Azure, GCP)

Finance Domain Knowledge:

- Financial markets and instruments
- Risk management principles
- Regulatory environment
- Business context

Soft Skills:

- Problem framing
- Critical thinking (avoiding overfitting)
- Communication (explaining models)
- Ethics and responsibility
- Collaboration (cross-functional teams)

Career Paths:

- Quantitative Analyst
- Data Scientist (Finance)
- ML Engineer
- Risk Modeler
- AI Product Manager

Key Insight: Success requires combination of technical skills, domain expertise, and ethical awareness.

Summary and Key Takeaways

Core Concepts:

- AI ¿ ML ¿ Deep Learning (hierarchy)
- ML learns patterns from data
- Supervised learning dominates finance
- Deep learning for unstructured data

Finance Applications:

- Risk: Credit, fraud, AML
- Trading: Algorithms, robo-advisors
- Operations: NLP, automation
- Customer: Chatbots, personalization

Reality Check:

- AI augments, not replaces humans
- Data quality is paramount
- Hype vs. realistic capabilities
- Ethical and regulatory challenges

Looking Ahead:

- Next 11 lessons: Detailed exploration
- Hands-on understanding (conceptual)
- Critical evaluation skills
- Practical applications

Next Lesson Preview

Lesson 26: Financial Data for AI/ML

Topics to be covered:

- Structured vs. unstructured data
- Data sources and vendors
- Alternative data revolution
- Data quality and preprocessing
- GDPR and privacy considerations
- Feature engineering basics

Preparation:

- Review basic statistics (mean, variance, correlation)
- Think about data quality issues in your own experience
- Consider: What makes financial data unique?