1. INTRODUCTION

Fintech (finance + technology) is playing a major role in the fields of:
- investment management industry
- investment advisory services
- financial record keeping, blockchain and distributed ledger technology (DLT)

2. WHAT IS FINTECH

Some salient fintech developments related to the investment industry include:
- **Analysis of large data sets:**
  - traditional data sources include economic indicators, financial statements
  - non-traditional data sources (such as social media, sensor networks) to generate profits.
- **Analytical tools:** artificial intelligence (AI) helps identifying complex, non-linear relationships among gigantic datasets.
- **Automated trading:** lower transaction costs, market liquidity, secrecy, efficient trading etc.
- **Automated advice:** Robo-advisors or automated personal wealth management are low-cost alternates for retail investors.
- **Financial record keeping:** DLT provides advanced and secure means of record keeping and tracing ownership of financial assets on peer-to-peer (P2P) basis.

3. BIG DATA

Big data typically have the following features:
- Volume
- Velocity
- Variety

<table>
<thead>
<tr>
<th>Traditional</th>
<th>Non-traditional (alternate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sources</td>
<td>Sources</td>
</tr>
<tr>
<td>Institutions, Businesses, Government, Financial Markets</td>
<td>Social media, Sensor networks Company-used data, Electronic devices, Smart phones, Cameras, Microphones, Radio-frequency identification (RFID)</td>
</tr>
<tr>
<td>Forms of Data</td>
<td>Forms of Data</td>
</tr>
<tr>
<td>Annual reports, Regulatory filings, Sales &amp; earnings, Conference calls, Trade prices &amp; volumes</td>
<td>Posts, Tweets, Blogs, Email, Text messages, Web-traffic, Online news sites</td>
</tr>
</tbody>
</table>
3.1 Sources of Big Data

Main sources of alternative data are data generated by:
1. **Individuals**: Data in the form of text, video, photo, audio or other online activities (customer reviews, e-commerce).
2. **Business processes**: data generated by corporations or other public entities e.g. sales information, corporate exhaust.
3. **Sensors**: data connected to devices via wireless networks.

3.2 Big Data Challenges

In investment analysis, using big data is challenging in terms of its:
- quality (selection bias, missing data, outliers)
- volume (data sufficiency)
- suitability

4. ADVANCED ANALYTICAL TOOLS: ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

**Artificial intelligence (AI)** technology in computer systems is used to perform tasks that involve cognitive and decision-making ability similar or superior to human brains.

**Machine learning (ML)** algorithms are computer programs that perform tasks and improve their performance over time with experience.

ML divides data into two sets:
- **Training data**: that helps ML to identify relationships between inputs and outputs through historical patterns.
- **Validation data**: that validates the performance of the model by testing the relationships developed (using the training data).

4.1 Types of Machine Learning

Two main types of machine learning are:
- **Supervised learning**: uses labeled training data and process that information to find the output. Supervised learning follows the logic of 'X leads to Y'.
- **Unsupervised learning**: does not make use of labelled training data and does not follow the logic of 'X leads to Y'. There are no outcomes to match to, however, the input data is analyzed, and the program discovers structures within the data itself.

**Deep Learning Nets (DLNs)**: Some approaches use both supervised and unsupervised ML techniques. DLNs use neural networks often with many hidden layers to perform non-linear data processing.
5. DATA SCIENCE: EXTRACTING INFORMATION FROM BIG DATA

5.1 Data Processing Methods

Five data processing methods are:

i) Capture: Data capture refers to how data is collected and formatted for further analysis.

ii) Curation: Data curation refers to managing and cleaning data to ensure data quality.

iii) Storage: Data storage refers to archiving and storing data.

iv) Search: Search refers to how to locate requested data.

v) Transfer: Data transfer refers to how to move data from its storage location to the underlying analytical tool.

5.2 Data Visualization

Data visualization refers to how data will be formatted and displayed visually in graphical format. Data visualization for:

- **traditional structured** data can be done using tables, charts and trends.
- **non-traditional unstructured** data can be achieved using new visualization techniques such as:
  - interactive 3D graphics
  - visualization techniques using colors, shapes, sizes etc.
  - tag cloud,
  - mind map

6. SELECTED APPLICATIONS OF FINTECH TO INVESTMENT MANAGEMENT

6.1 Text Analytics and Natural Language Processing

**Text analytics** is a use of computer programs to retrieve and analyze information from large unstructured text or voice-based data sources.

**Natural language processing (NLP)** is a field of research that focuses on development of computer programs to interpret human language. NLP field exists at the intersection of computer science, AI, and linguistics.

6.2 Robo-Advisory Services

Robo-advisory services provide online programs for investment solutions without direct interaction with financial advisors.

**Two types of robo-advisory wealth management services are**:

- Fully Automated Digital Wealth Managers
- Advisor-Assisted Digital Wealth Managers

**Limitations of Robo-advisors**

- The role of robo-advisors dwindles in the time of crises.
- The rationale behind the advice of robo-advisors is not fully clear.
- The trust issues with robo-advisors may arise specially after they recommend some unsuitable investments.
- As the complexity & size of investor’s portfolio ↑, robo-advisor’s ability to deliver detailed & accurate services ↓.

6.3 Risk Analysis

Advanced AI techniques are helping managers in performing scenario analysis i.e. hypothetical stress scenario, historical stress event, what if analysis, portfolio backtesting etc.

**Stress testing and risk assessment measures** require wide range of quantitative and qualitative data.

Big data and ML techniques may provide intuition into real time to help recognize changing market conditions and trends in advance.

6.4 Algorithmic Trading

Computerized trading of financial instruments based on some pre-specified rules and guidelines.

**Benefits**:

- Execution speed
- Anonymity
- Lower transaction costs

**High-frequency trading (HFT)** is a kind of algorithmic trading that execute large number of orders in fractions of seconds.
Distributed ledger technology (DLT) – advancements in financial record keeping systems – offers efficient methods to generate, exchange and track ownership of financial assets on a peer-to-peer basis.

**DLT advantages:**
- i) accuracy
- ii) transparency
- iii) secure record keeping
- iv) speedy ownership transfer
- v) peer-to-peer interactions

**Limitations:**
- i) excessive energy consumption
- ii) not fully secure technology

Three basic elements of a DLT network are:
1. **Digital ledger** – a digital database to record & store transactions
2. **A consensus mechanism** - mechanism which ensures that entities verify the transactions and agree on the common state of the ledge
3. **Participant network** – a peer-to-peer network of nodes.

A **distributed ledger** is a digital database where transactions are recorded, stored and distributed among entities in a manner that each entity has a similar copy of digital data.

**Blockchain** is a digital ledger where transactions are recorded serially in blocks that are then joined using cryptography.

**7.1 Permission and Permissionless Networks**

**7.2 Application of Distributed Ledger Technology to Investment Management**

**7.2.1 Cryptocurrencies**

**7.2.2 Tokenization**

**7.2.3 Post-trade clearing and settlement**

**7.2.4 Compliance**

DLT networks can be permissionless or permissioned.

**Permissionless networks** are open to new users. Participants can see all transactions and can perform all network functions.

**Permissioned networks** are closed networks where activities of participants are well-defined. Only pre-approved participants are permitted to make changes.