Introduction to Machine Learning and Applications in FinTech

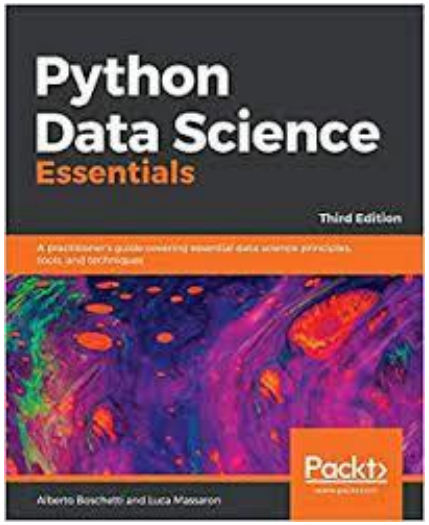
Lecture 1

Outline

- Textbook
- Course Contents
- Introduction of Machine Learning
- Applications in FinTech

Textbook

- Alberto Boschetti and Luca Massaron. *Python data science essentials: A practitioner's guide covering essential data science principles, tools, and techniques*. Packt Publishing Ltd, 2018.
- https://www.packtpub.com/product/python-data-science-essentials-third-edition/9781789537864
- We don't cover the whole book and technique details. Students only need to know the selected contents.



Course Contents I

1) Introduction to the Module, Machine Learning and Applications in FinTech

2) Data management in Python

3) Dimensionality reduction, EDA, and visualization

4) Supervised methods - A: Regression and Naive Bayesian **+ Coursework 1 Detailed Guidelines**

5) Supervised methods - B: K-nearest neighbor and Support vector machines

Course Contents II

6) Supervised methods - C: Decision trees

7,8) Supervised methods - D: Bagging and Boosting

9) Unsupervised methods: K-means clustering

+ Coursework 2 Detailed Guidelines

10) Introduction to deep learning and neural network

11) Review

Feedback and suggestions

- Your feedback and suggestions are welcome.
- Feel free to email me with suggestions at any time or let me know during all sessions.
- You can tell me any feedback and suggestions directly anytime.

Deep Learning

What society thinks I do

What my friends think I do

What other computer scientists think I do

What mathematicians think I do

What I think I do

```
In [1]:
import keras
Using TensorFlow backend.
```

What I actually do

Typical codes to run machine learning

```
from sklearn.linear_model import LinearRegression
hypothesis = LinearRegression()
hypothesis.fit(X_train, Y_train)
Y_pred = hypothesis.predict(X_test)
```

We DON'T Cover

Complicated source codes of Python libraries

```
for(i=0;i<l;i++)
{
	for(m=0;m<nr_class;m++)
		alpha_index[i*nr_class+m] = m;
	feature_node *xi = prob->x[i];
	QD[i] = 0;
	while(xi->index != -1)
	{
		double val = xi->value;
		QD[i] += val*val;

		// Uncomment the for loop if initial alpha isn't zero
		// for(m=0; m<nr_class; m++)
		//	w[(xi->index-1)*nr_class+m] += alpha[i*nr_class+m]*val;
		xi++;
	}
	active_size_i[i] = nr_class;
	y_index[i] = (int)prob->y[i];
	index[i] = i;
}
```

Complicated mathematics and theories of machine learning

1.4.7.3. SVR

Given training vectors $x_i \in \mathbb{R}^p$, i=1,..., n, and a vector $y \in \mathbb{R}^n$ ε-SVR solves the following primal problem:

$$\min_{w,b,\zeta,\zeta^*} \frac{1}{2} w^T w + C \sum_{i=1}^{n} (\zeta_i + \zeta_i^*)$$

$$\text{subject to } y_i - w^T \phi(x_i) - b \leq \varepsilon + \zeta_i,$$
$$w^T \phi(x_i) + b - y_i \leq \varepsilon + \zeta_i^*,$$
$$\zeta_i, \zeta_i^* \geq 0, i = 1, \ldots, n$$

Its dual is

$$\min_{\alpha,\alpha^*} \frac{1}{2} (\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + \varepsilon e^T (\alpha + \alpha^*) - y^T (\alpha - \alpha^*)$$

$$\text{subject to } e^T (\alpha - \alpha^*) = 0$$
$$0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \ldots, n$$

where e is the vector of all ones, $C > 0$ is the upper bound, Q is an n by n positive semidefinite matrix, $Q_{ij} \equiv K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel. Here training vectors are implicitly mapped into a higher (maybe infinite) dimensional space by the function ϕ.

The decision function is:

$$\sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) + \rho$$

https://github.com/scikit-learn/scikit-learn/blob/main/sklearn/svm/src/liblinear/linear.cpp

Methods for studying the module

This is a practical module of applying classic machine learning methods to finance cases and data. We will focus on repeatedly use a few easy lines of codes to different finance datasets.

This is **NOT** a mathematics module and **NOT** a computer science module. We will not discuss too much complex mathematics and codes that requires strong UG background in advance mathematics and coding.

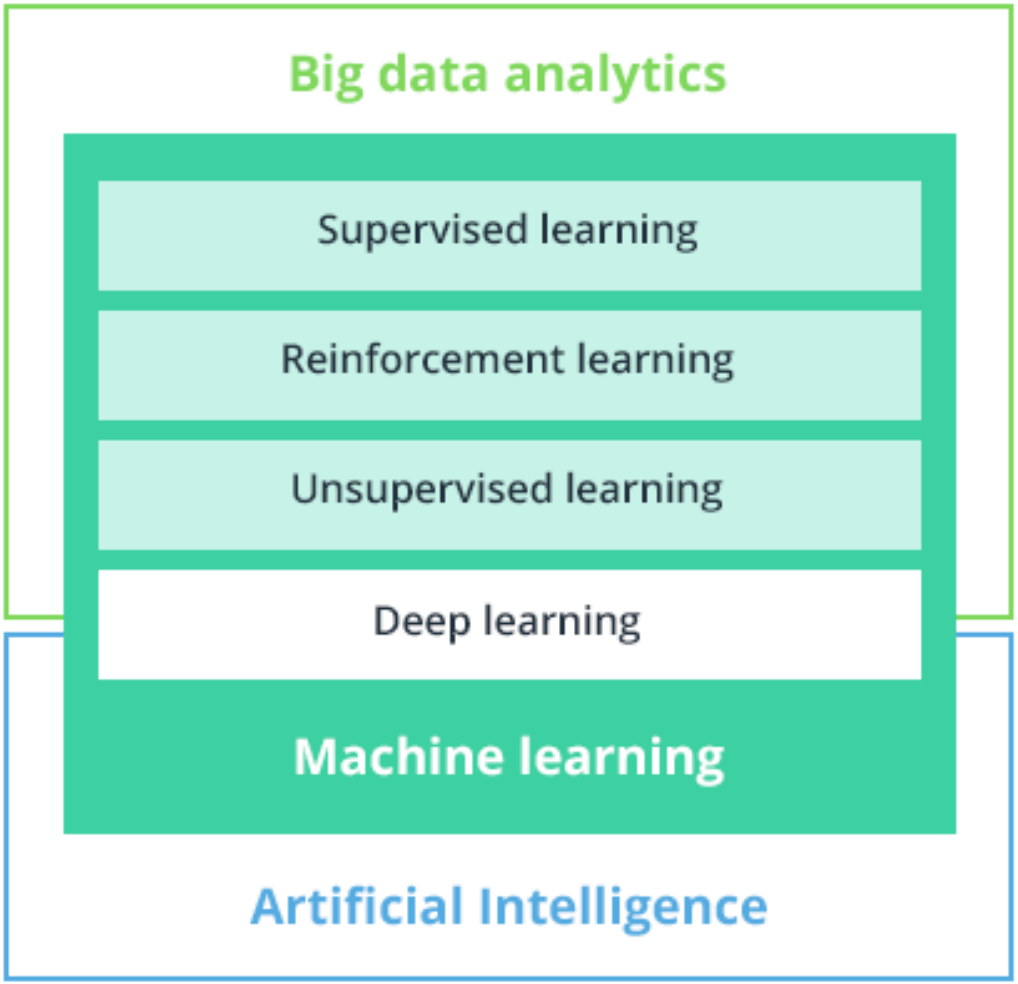
Learning data analytics is like learning language. We need to start from reading / learning examples and then we can use these examples for new problems.

Students will have a better understanding of methods gradually after studying intuition and a few easy lines of codes, which needs a learning process. following an organised structure for helping students understand.

In the first part of the module, we spend more time in building the foundation from the simple textbook examples.

In the second part, we look at more real applications outside of the textbook, such as the **case studies of trading strategies, IPO, UK firms' credit, and Lending Club**.

A schematic view of ML in relation to AI and big data analytics



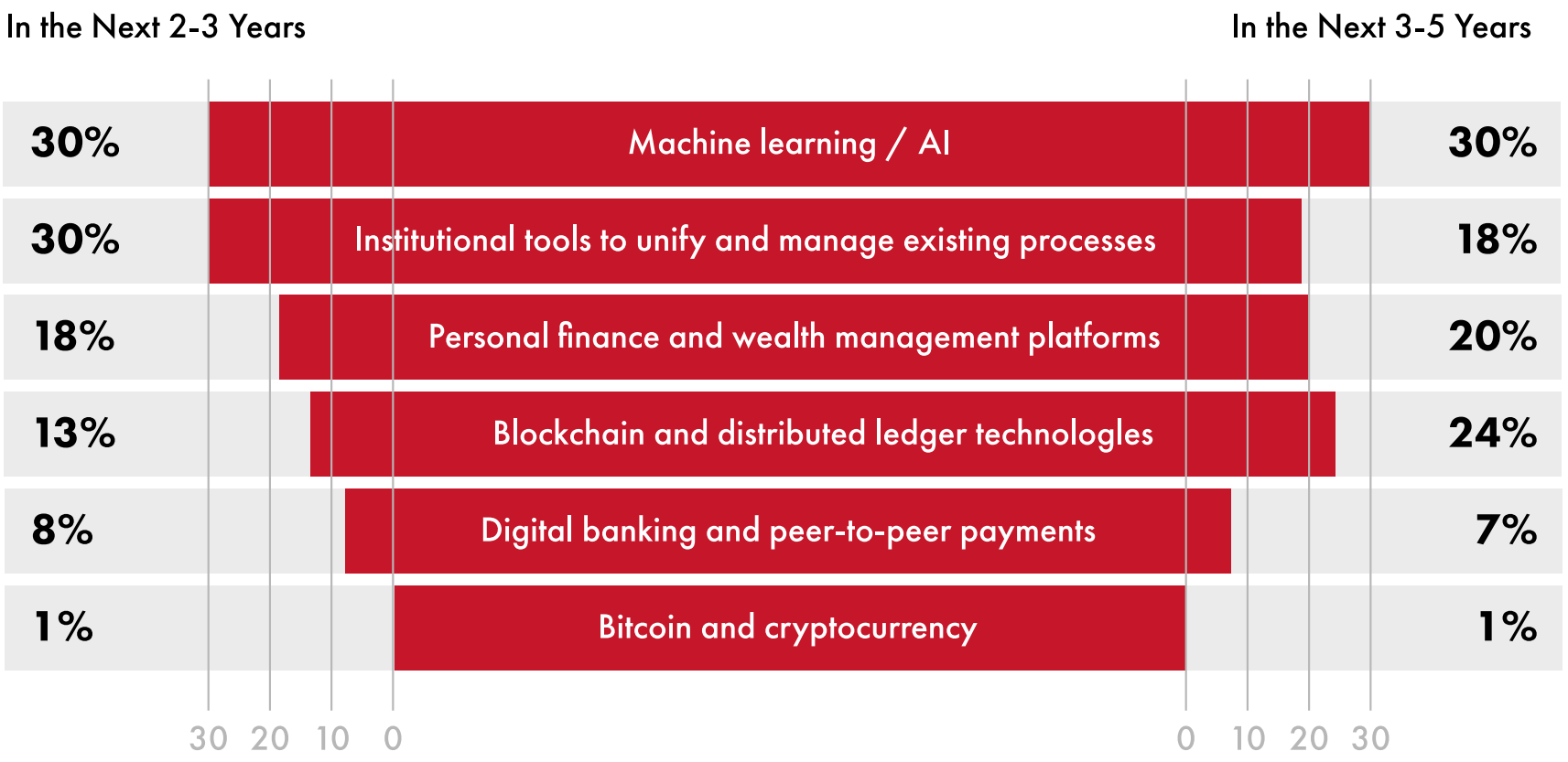
Source: The Financial Stability Board (FSB) – Artificial intelligence and machine learning in financial services

A typical workflow for a trading system using supervised learning

Source: A Machine Learning Framework for Algorithmic Trading on Energy Markets

What is Machine Learning in FinTech

- Learn and improve from experience and large volumes of data
- A key aspect of several financial services: managing assets, evaluating risk, calculating credit scores, detecting fraud, automate trading, and providing financial advisory etc.



Source: Mediant FinTech Trends Report

Forecast the stock market

- The vast volumes of trading operations result in tons of data — an unlimited potential for learning.
- ML algorithms monitor data sources available in real time, such as news and trade results, to pinpoint patterns indicating stock market dynamics.
- The task left to traders is to determine which ML algorithms to include in their strategies, make a trading forecast, and choose a behavioural pattern.

- The predictive capacities of machines are large.
- Machine learning can detect the slightest indicators of prices going up or down.
- Machine learning can easily compare data over several decades.
- Machine learning algorithms can make trading decisions extremely quickly.
- No bias from human interpretation.

- Sentient Technologies is an AI company that's developing and applying proprietary quantitative trading and investment strategies using distributed artificial intelligence systems.

- Walnut Algorithms is a start-up that offers AI and ML finance solutions for investment management.

- I Know First is a company offering stock forecasts based on predictions of machine learning algorithms.

Algorithmic trading

- Use algorithms to make better trade decisions.
- Build models with parameters that monitor news and trade activities to detect factors that forecast security prices.
- Simultaneously analyse large volumes of data and make thousands of fast trades every day.
- Judgment is not affected by emotions.
- Widely adopted by hedge fund managers.

Fraud detection and prevention

- Fraud is a high threat and accounts for billions of dollars in losses each year.
- Large amount of finance data stored online and it increases the risk of a security breach.
- Machine learning scans through large data sets to detect anomalies for further investigation.
- Compare a transaction against other data points, e.g., account history, IP address, location, etc.

Source: Maruti Techlabs – How Machine Learning Facilitates Fraud Detection

Rule-based Fraud Detection	ML-based Fraud Detection
• Catching obvious fraudulent scenarios	• Finding hidden and implicit correlations in data
• Requires much manual work to enumerate all possible detection rules	• Automatic detection of fraud scenario
• Multiple verification steps that harm user experience	• The reduced number of verification measures
• Long-term processing	• Real time processing

Source: Intellias

- Feedzai is a start-up that offers one of the most mature machine learning engines, which is quick at taking advanced fraud prevention measures.

- Biocatch combines behavioural biometrics with machine learning to recognize and prevent human and non-human cybersecurity threats mainly in banking, payments, and insurance.

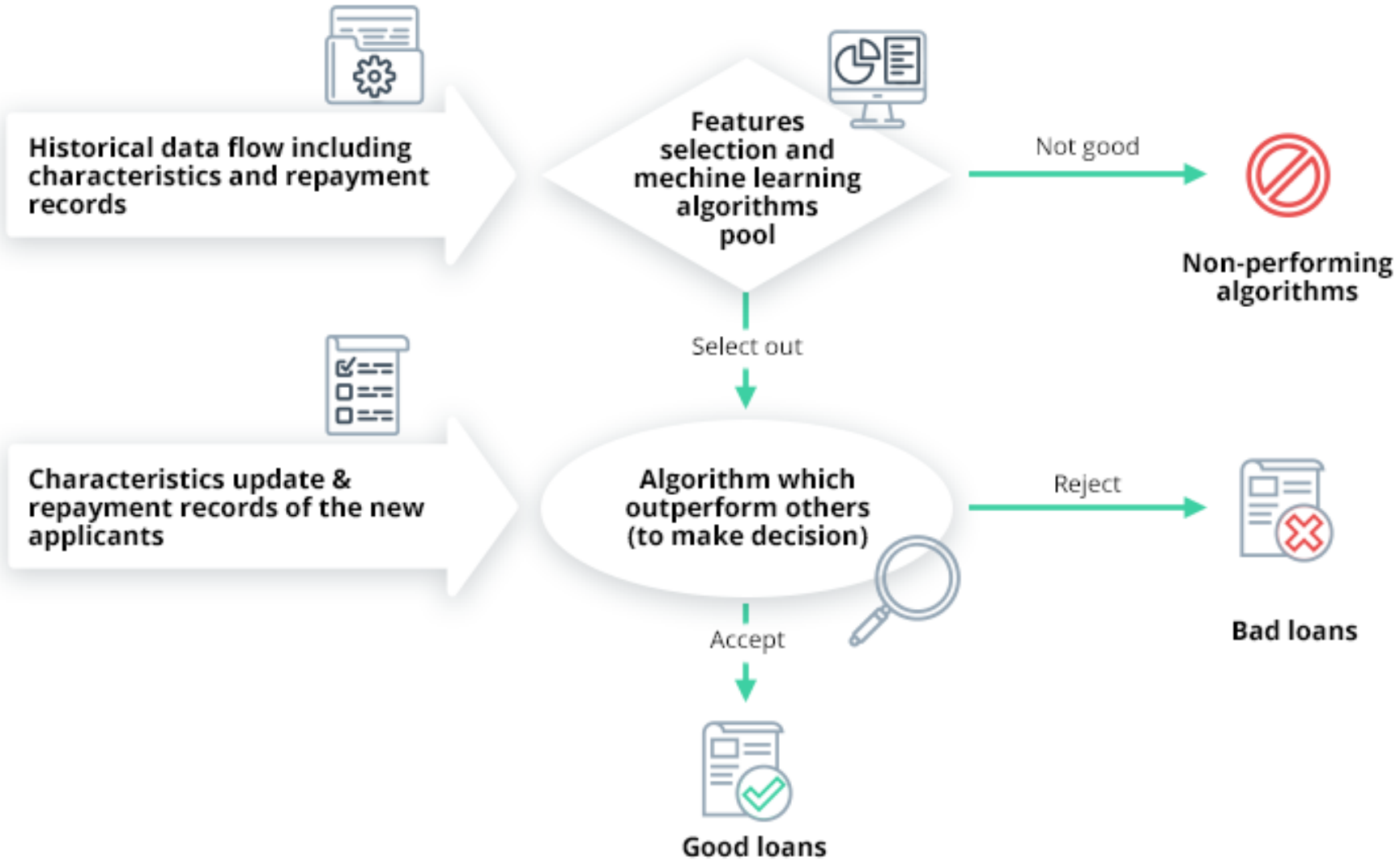
- Ravelin is a London-based company that uses machine learning to prevent and stop fraud in online payments.

Robo-advisors

- Use algorithms to establish a financial portfolio across different financial instruments and asset classes – such as stocks, bonds, real estate, etc. according to an investor's goals and risk tolerance.
- Low account minimums and are usually cheaper .
- Automatically determine the best investment opportunities according to real-time market trends to find the best diversification strategy.

Loan underwriting

- Machine learning can be trained with millions of consumer data in order to simplify the loan underwriting process and make quick decisions.
- Train algorithms to analyse millions of consumer data and look for unique exceptions.
- Consumer data: age, income, occupation, and the consumer's credit history, etc.
- Save companies both time and financial resources.



Source: Tieto – How machine learning can improve accuracy in credit scoring

- ZestFinance works on machine learning-based credit models to generate more profitable underwriting.

- Deserve uses machine learning to provide users with credit cards even if they have no credit score or need to rebuild their credit.

- Intellias assists lending providers with developing an ML-enabled credit score calculator. It runs with the help of ML algorithms to get data about borrowers and their businesses.

Regulatory compliance

- Billions of dollars spent on regulatory compliance were not enough for financial firms.
- Regulatory Technology (RegTech): ML algorithms can learn from a pile of regulatory documents.
- ML can detect correlations between guidelines.
- Automatically track and monitor regulatory changes as they appear.

Source: Intellias

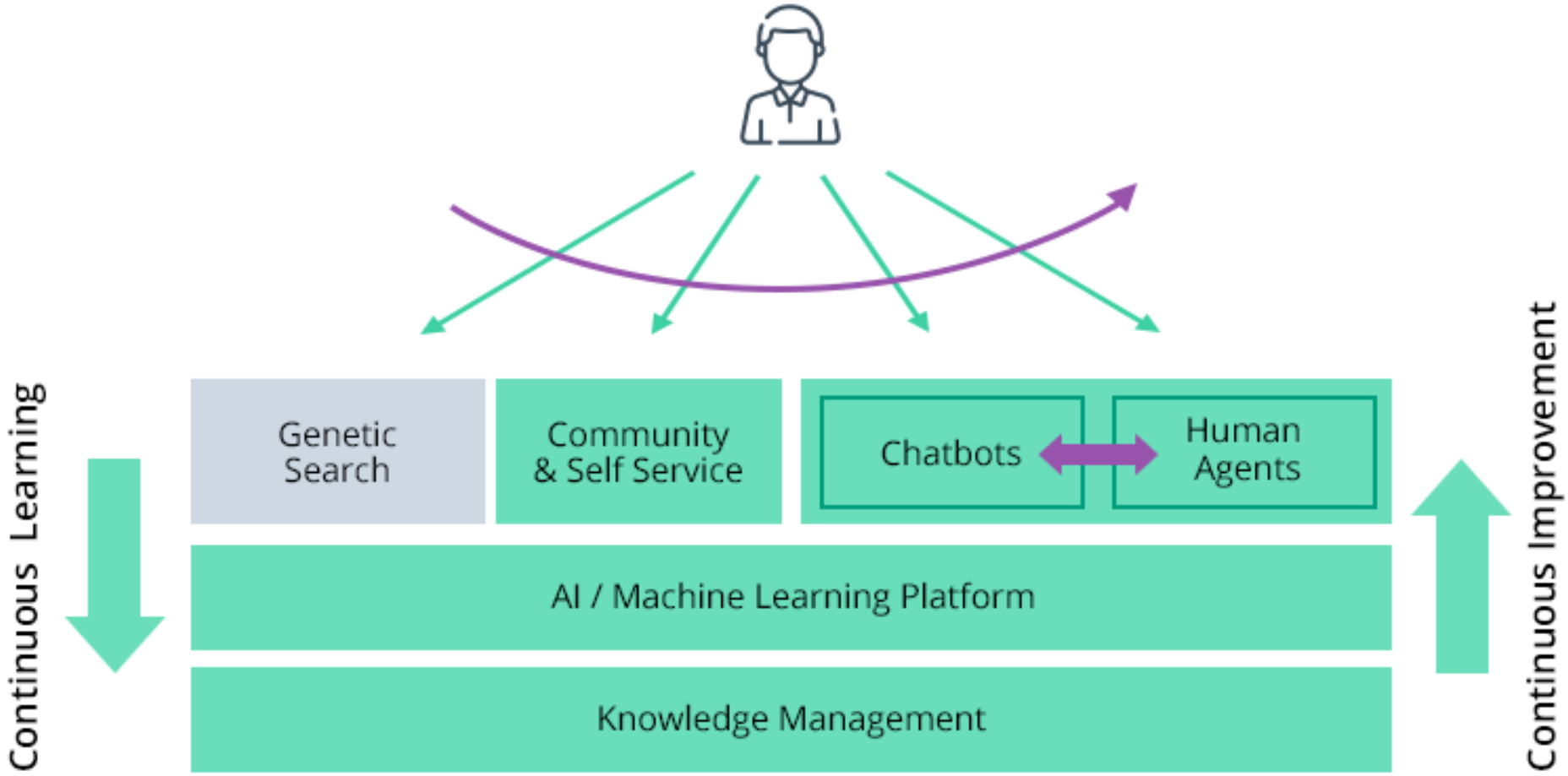
- Pendo Systems is a FinTech company that works with unstructured data to streamline the compliance process for their clients.

- Compliance.ai is a start-up that uses adaptive machine learning models in FinTech to automate research and track financial regulatory contents and regulatory updates in a single platform.

- ComplyAdvantage is a start-up that uses machine learning to accelerate FinTech compliance and enable online fraud prevention tools.

Automated Customer Support

- The need for better, safer, and customized solutions is rising with expectations of customers.
- Automation has helped the FinTech industry to provide better customer service and experience.
- AI interfaces and Chatbots can offer useful advice while reducing the cost of staffing.
- AI can automate the back office process and reduce human errors and improve customer support.



Source: CustomerThink – Customer Service beyond the Chatbot Hype

- Kasisto uses AI and ML algorithms to power omnichannel virtual assistants.

- Wells Fargo was the first US bank to launch an AI-driven customer chat experience for Facebook Messenger.

- Bank of America's Erica, an AI-based virtual assistant, was launched in March 2018 and helped more than 1 million users in the first three months.

Background scanning

- ML can significantly reduce the effort in the screening of the company before the transaction or any credit-related decisions.
- **Natural Language Processing** delivers sentiment analysis and check the contracts or emails.
- **Image Recognition** performed on satellite images can analyse the placement of shops or the routes that are used to transport goods.

FinTech Cases

- Being fully digital Banks, e.g., Monzo, enables customers to manage their financial lives directly from their smartphone.
- It invites customers to open an account in minutes, right from the smartphone. Customers manage it on-the-go, spend and set aside money in real-time.
- Previously complex processes can happen anywhere in minutes.

FinTech Cases

- BlackRock Investment Company, offers Aladdin, an operating system created and adapted for investment managers' needs of risk analytics and portfolio management software tools to make more informed investment decisions and operate more efficiently.

FinTech Cases

- Citi's and Trade Solution Group with Feedzai, a leader for real-time risk management
- Provide clients with enhanced control and risk management for payments transactions.
- The ML solution compares all the possible data points in current and preceding transactions to detect suspicious transactions with compliance concerns.

FinTech Cases

- Any large telecommunications company or mobile operator, e.g., Netflix, Amazon, Apple, TikTok.
- Forecasting customer preference,
- Detecting potential or current subscription customers,
- Sending commercial mailings with tailoring various loyalty programs for customers.

References

- Intellias.com, 5 Use Cases of Machine Learning in Finance and Banking.
- Corporatefinanceinstitute.com, Machine Learning in Finance.
- Fayrix.com, Best Use Cases of Machine Learning in Finance.
- Fintechnews.org, How AI amp Machine Learning is infiltrating the Fintech industry.
- Ideamotive.co, AI and Machine Learning Applications In Finance FinTech And Banking.
- Koombea.com, 8 Uses of AI in FinTech Using Machine Learning