Application of ML industries 1

Computer Science (Anna University)

Studocu is not sponsored or endorsed by any college or university
Downloaded by Rohit Srivastava (rts080185@gmail.com)

Smarter education for smarter planet

IBM Innovation Centre for Education

Welcome to:

Machines Learning in Banking and Securities

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Unit objectives

After completing this unit, you should be able to:

- Learn about machine learning in banking sectors, challenges in banking sector and fraud detection system in banking sector
- Understand manage customer data and algorithms in banking and security
- Gain knowledge on deep learning technology for personalized marketing
- Learn about applications of machine learning classifiers in credit risk analysis
- Understand the importance of machine learning in portfolio management systems
- Learn about algorithmic trading and stages for implementing applications of ai in marketing
- Deep learning approach for sentiment analysis of customers and customer services
- Understand cyber security in banking sector and loan underwriting and sentiment/news analysis
- Understand current challenges and opportunities in implementing machine learning technologies in banking and security domain

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Couse description

- Purpose: Machine learning in banking and securities.
- Audience: 4th sem engineering students.
- Perquisite: Probability, Python basics, statistics, linear algebra.
- Course objectives: To introduce participants to understand the importance machine learning applications to address the problems and challenges in machine learning and securities.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Why machine learning in banking sector

- After financial crisis around the world, there has been a significant transformation in the banking services across the globe with intervention of Artificial Intelligence (AI) applications.

- With appropriate and robust machine learning algorithms, there is a lot of potential to address the key issues in the banking sector, leading to digital transformation and enhanced services.

- Major areas of banking with potential AI intervention include Anti Money Laundering(AML), Chabot's, fraud detection, algorithmic trading, and digitization.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Use of AI in banking and finance

- Efficient resource management, improvement in performance of operations and smart decision making are few of the promising advantages of machine learning technologies in the banking and finance sector.

- Fraud detection, automated customer support and services, banking security are few of the important use cases of ML in banking sector.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Challenges in banking sectors and securities

- Fraud detection.
- Though competitions.
- Risk modelling.
- Customer data management.
- Decreased customer experience and loyalty.
- Personalized marketing.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Fraud detection (1 of 2)

- One of the foremost requirements in banking sector is to provide enough security to the customers and employees in terms of operations and financial transactions.

- MI techniques can help banks to detect fraud and can restrict account activities based on authentication and validation.

- Fraud detection includes the following initial steps:
  - Preliminary data testing and data sampling for model estimation.
  - Creating/estimating the appropriate model.
  - Deployment of the model and testing.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Fraud detection (2 of 2)

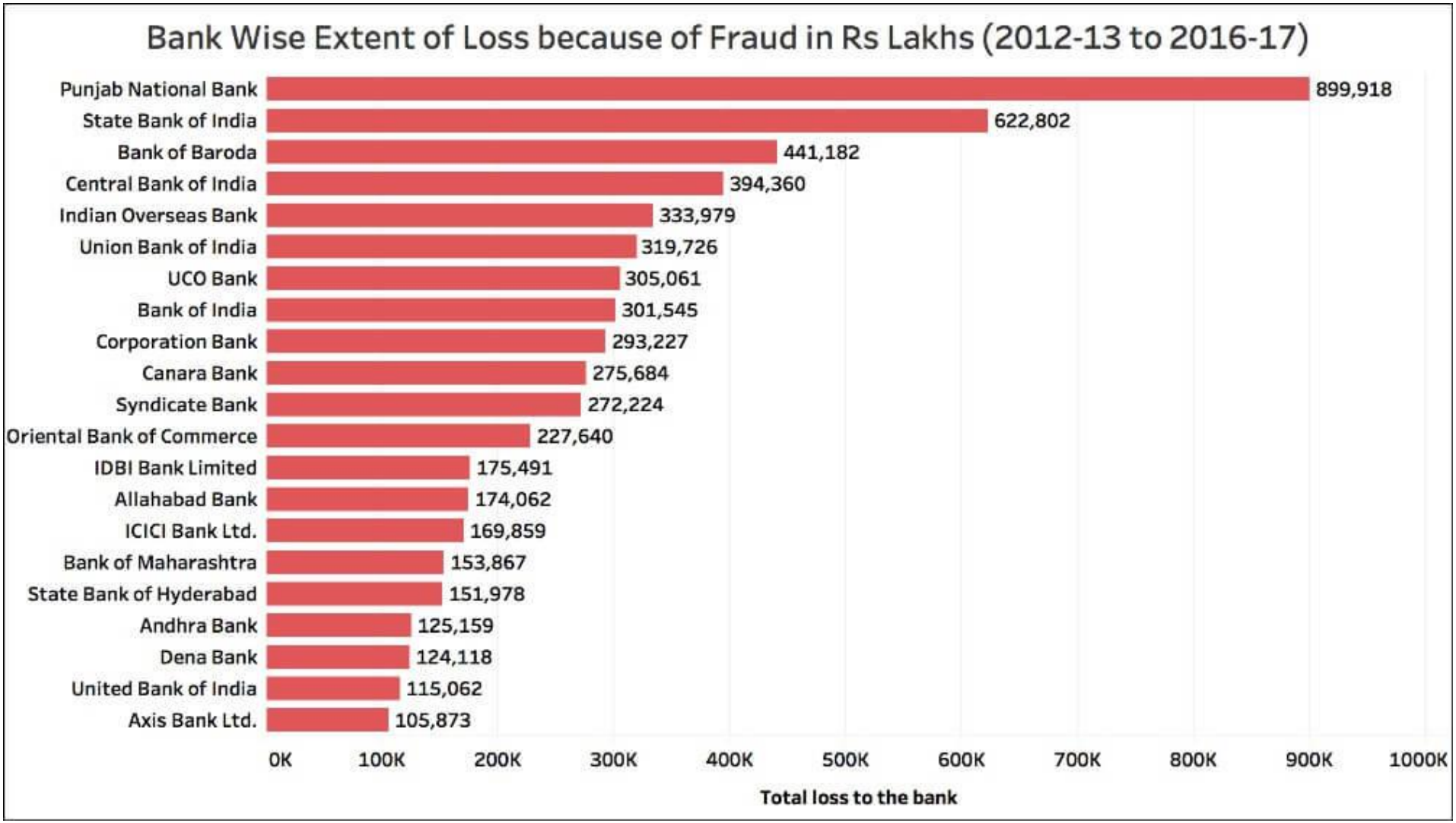


Figure: Fraud detection

Source: Factly

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Tough competition in banking industry

- Indian banking services has evolved with many “pro-customer” applications leading to competition among major retail banks.

- User friendly online banking applications like Paytm, Phone pay, Google pay are redefining the banking services, especially in terms of its geographically limitless reachability.

- Personalized banking experience is another most important aspect of the afore mentioned applications.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Risk modeling and investment banks

- The foremost priorities for any investment banks is managing financial risks.

- With effective risk modelling, it becomes easier for banks to evaluate and strengthen the capabilities of companies in creating financial capital, sustain or restructure corporate operations, facilitate acquisitions and mergers whenever appropriate.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Customer data management

- Financial organizations have a challenging task of documenting and monitoring humongous amount of customer data.
- It is essential to establish cross data set linkages to develop unique insight of customer needs.
- Data science can help banks extract relevant information about their customer behavior, whereas appropriate ML algorithms can keep track of customer preferences.
- One such example for customer data management is “segmenting the customers”. It refers to clustering of customers based on their unique behavior.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Decreased customer experience and loyalty

- Ever increasing customer focused services provided by companies like phone pay, google pay etc., is redefining the customer experience.

- Now a day, customers expect totally hustle free and easy banking services and it is more likely that banks can lose customers if they can not provide quality services to its customers.

- Banks can make use of efficient emotion recognition algorithms to deeply analyze the customer feelings and sentiments.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Personalized marketing

- The key aspect that leads to success in marketing is, the extent to which customers receives offers that best suits their needs and preferences.

- In the banking sector, the financial products should be highly customized to the customers' needs and in addition, secured as well.

- As the customers are exposed to such personized banking experiences, there will be ever increasing expectation for improvement in the quality of services.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Role of machine learning: Challenges of banking sector and securities

Domain	Implications of ML based solutions
Fraud and Risk management	ML based solutions and predictive analytics are assisting in examination of real-time transactions to identify suspicious and fraudulent operations. Risk analysis experts are being guided by ML algorithms with appropriate recommendations to predict risk in the earlier stages of any banking operations.
Customer Services	With existing customer data along with ML powered AI applications are leading to effective personalized customer services by documenting and analyzing customer behavior and requirements. ML algorithms based cognitive machines are replacing humans in analyzing and responding to customer queries.
Financial trading and securities	ML based validation mechanisms are bridging the security and functional gap between front end trades and back end operations. AI based applications are assisting banks in effectively handling foreign exchange transactions and liquidity management operations.
Credit assessment	ML based applications along with big data analytics are viable solution to assess the credit worthiness of the customer in case of loan disbursement operations.
Portfolio management	AI and ML based technological ecosystems are helping banks in making real time, smarter decisions to ensure appropriate investment plans for their customers.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Widely used machine learning algorithm in banking and security

- Supervised machine learning algorithms.
- Unsupervised machine learning algorithms.
- Reinforcement algorithms.

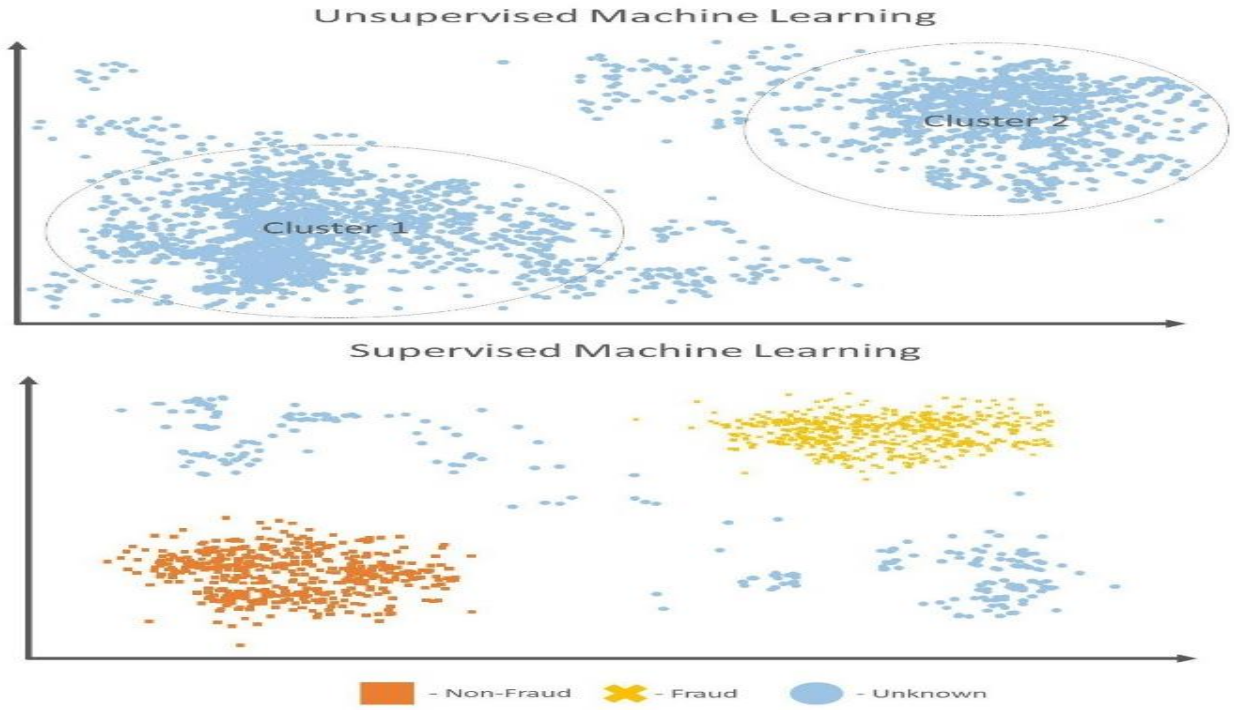


Figure: Machine learning algorithms in Banking and security

Source: Altexsoft

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Fraud prevention and detection systems

- The most used algorithms in fraud prevention and detection systems are:
  - Bayesian algorithms.
  - K-Nearest neighbor.
  - Support Vector machines (SVM).
  - Bagging ensemble classifier based on decision tree.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Rule based and machine learning based approach in fraud detection

IBM ICE (Innovation Centre for Education)

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	Possible fraud scenario detection happens automatically
Multiple verification steps may become threat to user experience.	Reduction in the number of verification measures
Long term processing	Real time processing

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Anomaly detection: Ways to expose suspicious transactions in banks

- The main feature of this approach is classification of all data objects into two major groups:
  - Normal distribution.
  - Outliers.

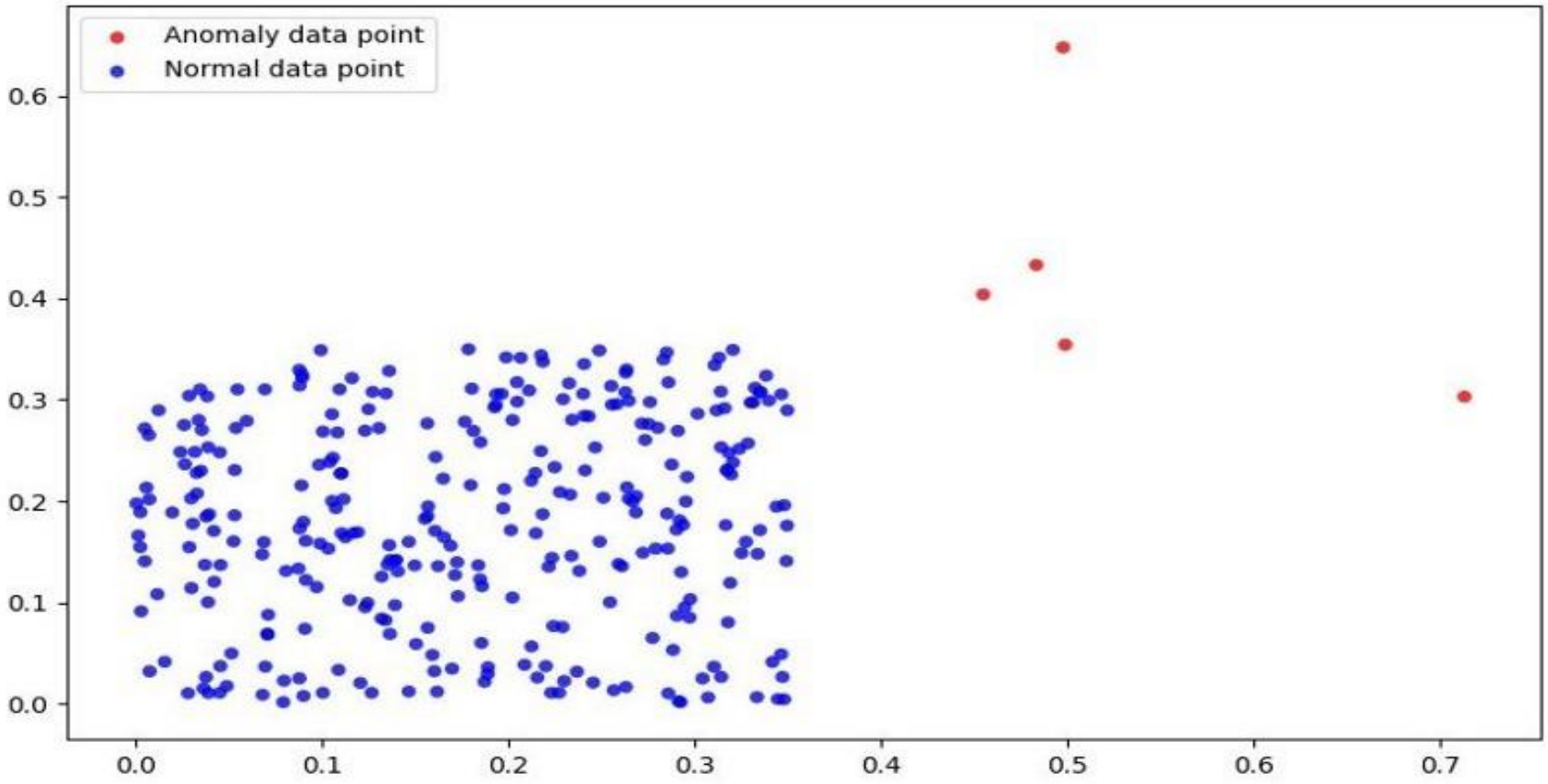


Figure: Ways to expose suspicious transactions in banks

Source: https://zindi.africa/blog/introduction-to-anomaly-detection-using-machine-learning-with-a-case-study

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Advanced fraud detection systems

- One of the major limitations of the primitive fraud detection approaches is that they are limited to identification of anomalies.

- The two commonly used ML techniques to develop anti-fraud mechanisms are:
  - Supervised ML algorithms.
  - Unsupervised ML algorithms.
  - Deep neural network algorithms.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Risk management systems

SL NO	Objectives of Portfolio Management	Key Features/ Objectives
1.	Security of Principal Investment	• To ensure safety investment and risk minimization • Along with keeping the investment unharmful, portfolio management also contribute in increase of its purchasing power over a period. • Only after ensuring safe investment, other factors like growth, income etc., are considered
2.	Consistency of returns	• In order to ensure stable return on investment by reinvesting the earned profit in good portfolios. • It helps the customers/ firms to obtain consistent returns.
3.	Capital growth	• Portfolio management helps the firms by increase in the rate of return by reinvesting securely or purchasing the profitable growth securities. • It takes care of any loss in purchasing power because of other economic factors, in order to appreciate values and to safeguard investors.
4.	Marketability	• In order to ensure flexibility in investment portfolio. • These portfolios consist of investment which concentrates only on the portfolios that can be marketed and traded. • Shifting from one investment to other will be a problem if the portfolio contains large number of inactive and unlisted shares. • It suggests investing only in secured stock exchanges.
5.	Liquidity	• It facilitates investors to make best use of available opportunities in future. • Based on the investor's requirement, it always ensure a enough fund availability in short notice
6.	Diversification of Portfolio	• To effectively manage the financial investment made by customer it is necessary to diversify the investment by considering benefits available across the industries.
7.	Favorable Tax status	• In order to minimize the tax burden and to increase the return in investors fund, the portfolio should assess by considering the applicable taxes.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

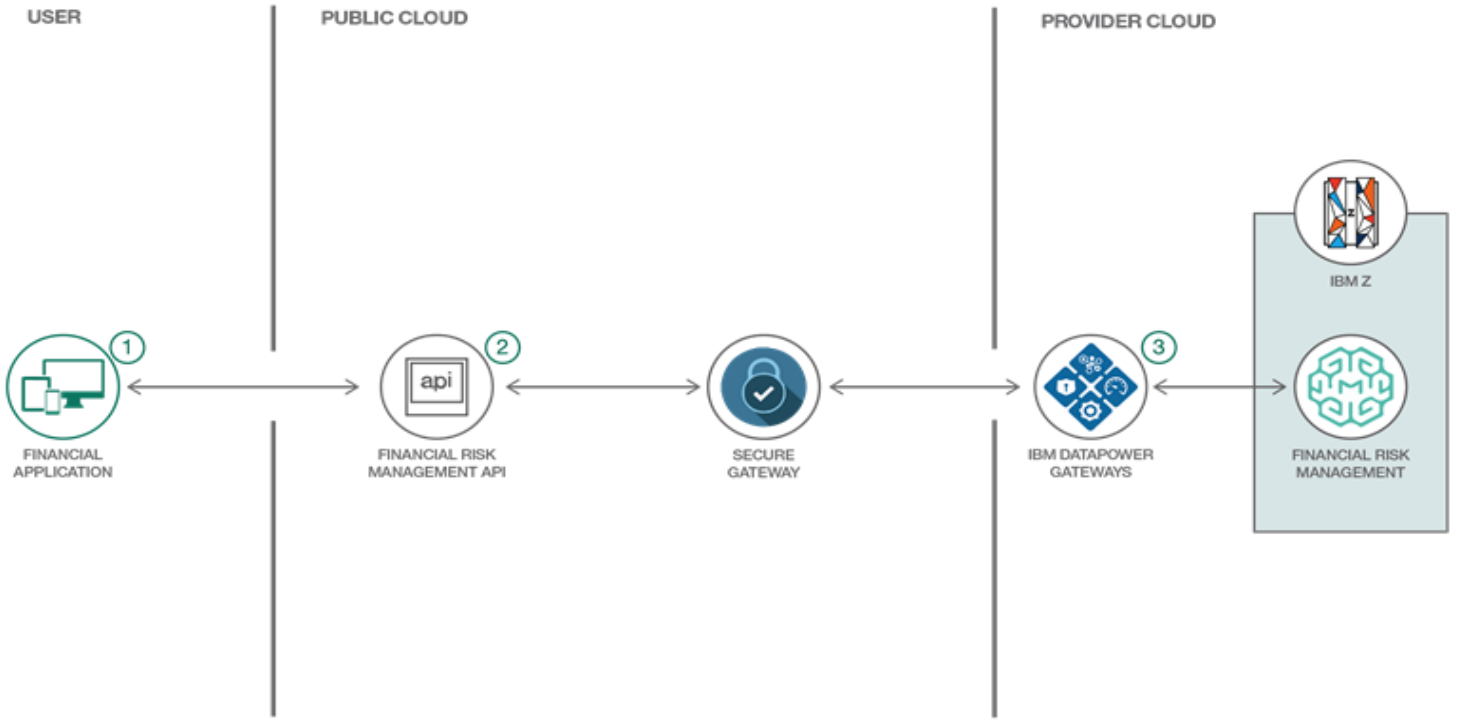
Case study: Application of machine learning to financial risk management

IBM ICE (Innovation Centre for Education)

- Summary: Machine learning affects all industry fields, like how banking companies and other sectors tackle stricter enforcement and threat mitigation criteria.

- This programmer path teaches you what to implement machine learning for a strategic threat system on IBM z/OS to assess the value of consumer debt. utilizing an API, you can know how to view the information, allowing you to integrate the information into company programs.

- Flow :



This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Credit risk analysis using machine learning classifier

- Promising business activity in banking industry is providing loan to the customers.
- Crediting risk analysis is evolving in the field of financial risk management.
- Based on the available database of customers many techniques were used in credit risk analysis in order to evaluate future credit risk that may occur.
- ML techniques can be used to analyze and evaluate the credit risk datasets.
  - Bayesian classifier.
  - Naive–Bayes classifier.
  - K-nearest neighbors.
  - K-means.
  - Multilayer perceptron.
  - Support vector machine.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Investment prediction systems

- The wealth management organizations are evolving with the potential AI based solutions for their investment decisions based on the historical data.

- The applications of collection of huge data about the assets are being recorded for digital assets or distributed industrial asset by making the assets ready for digitalization through AI.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Portfolio management systems

- Portfolio: It can be defined as financial assets like stocks, bonds, shares, mutual funds, cash equivalents, debt instruments etc., In order to maintain the risk in various asset pools of investment, a portfolio is planned.

- Management: In order to achieve its pre described objectives with well-defined policies, a management can be defined as a firm which coordinates the activities of that firm/enterprise.

No.	Investor's Portfolio	Investment	Percentage	Security	Returns
1.	Government Bonds	$ 25,000	25 %	High	Low
2.	Bank's Fixed Deposits	$ 15,000	15 %	High	Average
3.	Shares	$ 35,000	35 %	Low	High
4.	Mutual Funds	$ 25,000	25 %	Average	Average

- Note: This is just an example and may be taken as a slandered for the portfolio management.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Objectives of portfolio management

Figure: Objectives of portfolio management

Source: Image credits © moon rodriguez

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Algorithmic trading

- Algorithm trading enables to produce profit at a very high intensity which is difficult by individual buyers.
- Algebraic equation is an example of algorithm, with prescribed rules of algebra.
- Complex formulas along with mathematical models and human inaccuracy will be used by algorithmic model in order to sell or buy the financial securities.
- The use of high-frequency algorithmic trading technology, enables algorithmic traders to make 10000 trades in fraction of time.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Deep learning for customer services

- Customer service center, are ready to adopt technologies like ML for their operations and these techniques is part of industries in upcoming days.

- In order to provide efficient customer service, it is always more flexible if we gather the data from their insights.

- Effective machine learning systems are implemented in environments where a lot of information is stored, as is important when the ultimate objective is to make an informed judgment.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Chatbot-deep learning approach

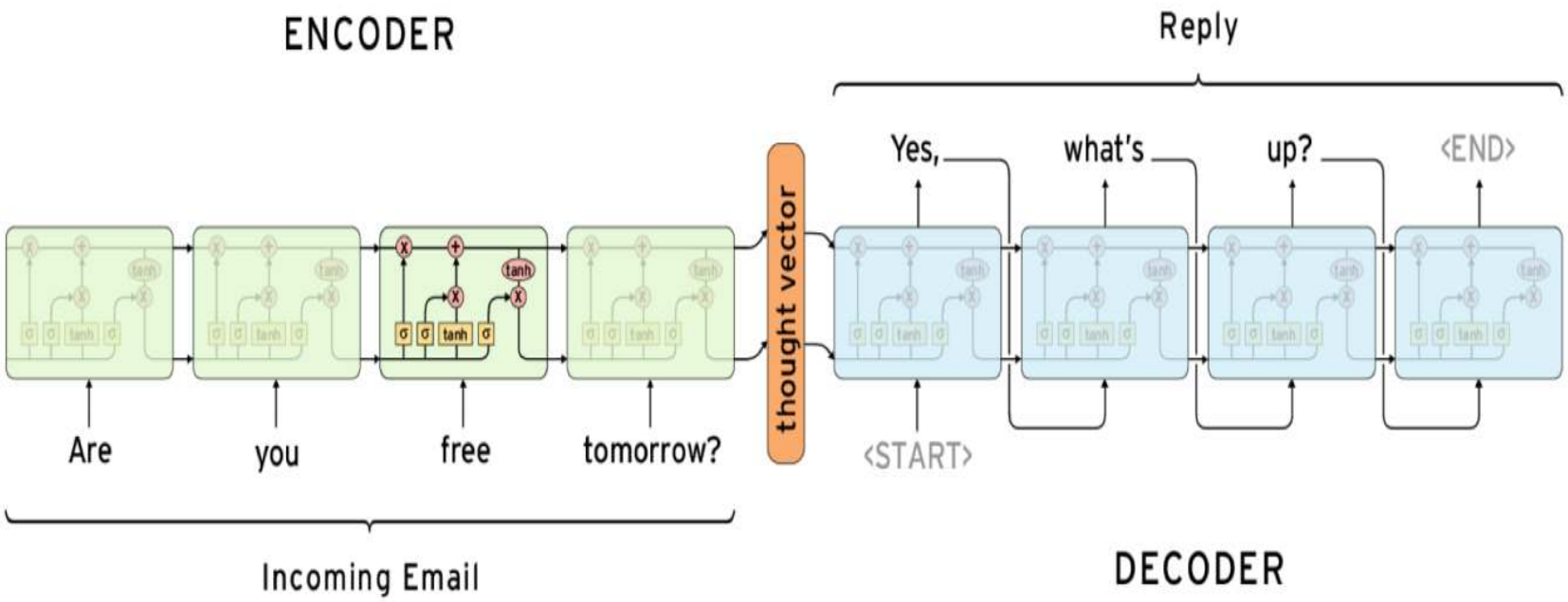


Figure: Chatbot-deep learning approach

Source: https://research.googleblog.com/2015/11/computer-respond-to-this-email.html

Downloaded by Rohit Srivastava (rts080185@gmail.com)

AI powered marketing systems

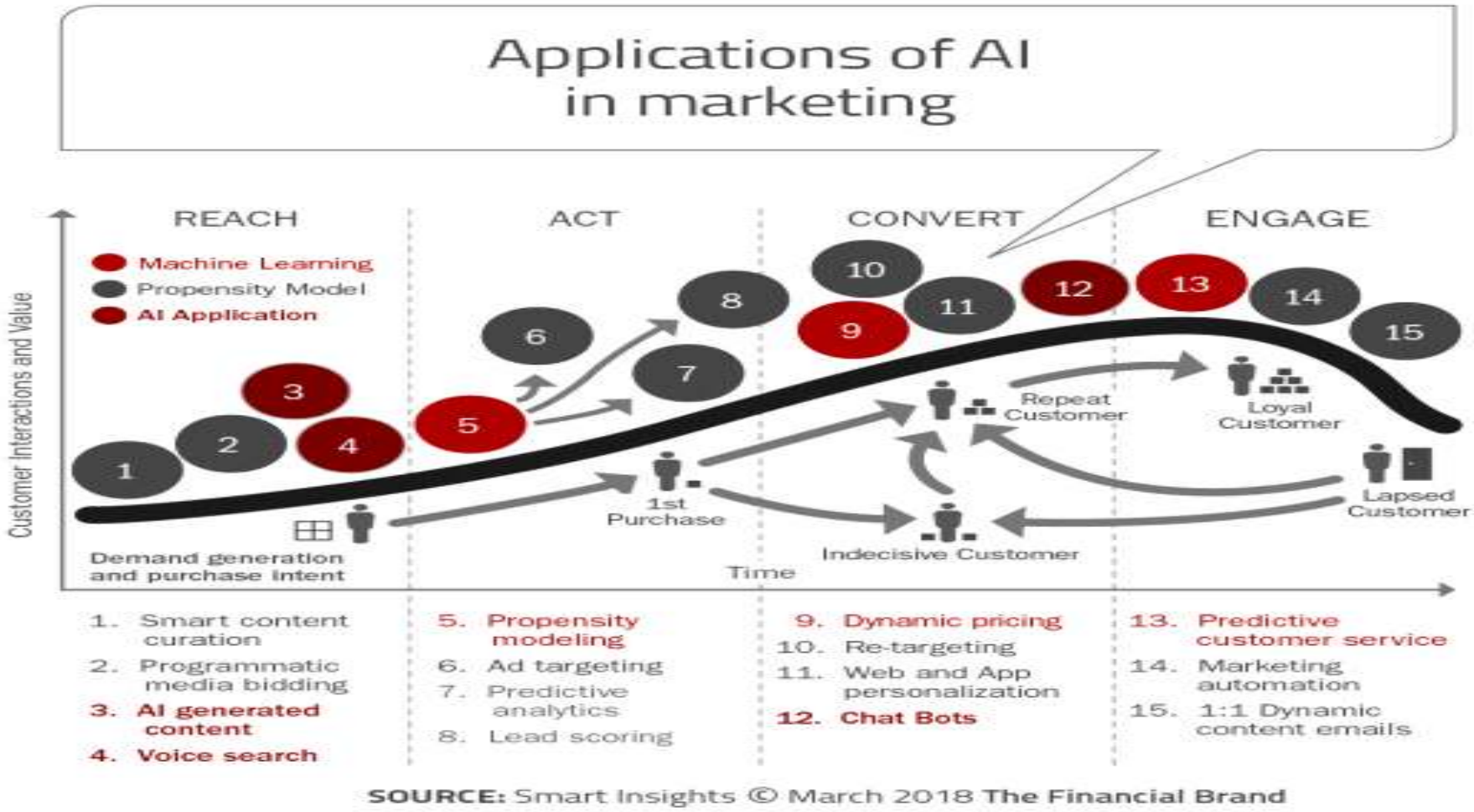


Figure: Stages for implementing applications of AI in marketing

Source: Smart insights © March 2018 the financial brand

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Deep learning in cyber security

- Today cyber security uses smart technology such as machine learning and natural language analysis that can enable security analysts make smarter, quicker judgments.

- For the hackers banking sector is the hub where they can grab those important information of the customer.

- Security and its team are facing challenges to handle, interpret and prevent these type of mischievous events.

- Having a robust cyber security is an important requirement to maintain strong customer trust and credibility.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Types of cyber–attacks in banks

IBM ICE (Innovation Centre for Education)

Denial of Services (DoS)	Phishing	Malware	Watering Hole	Zero-day exploits
Denial-of-service (DoS) attacks inundate systems with traffic to consume resources and bandwidth and make them unable to perform.	Phishing typically uses email that appears to be from a trusted or reputable source. Unsuspecting users open the email and take further actions like providing protected information or downloading malware.	Malware is malicious software. It's the chief weapon of a cyber-attack and includes viruses, worms, Trojans, ransomware, adware, spyware bots, bugs and rootkits. It installs when a user clicks a link or takes an action. When inside, malware can block access to data and programs, steal information and make systems inoperable.	In recent years' financial companies are most affected by watering hole, it is the most implemented cyber-attack. It mainly effects the IT system through online searches to figure out the behavioral pattern of employees to identify the website they frequently visit.	Zero-day exploits introduce malware through vulnerabilities unknown to the maker or user of software or systems. It is "zero-day" because developers have had zero-time to address or patch the vulnerability.[2]

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Deep learning methods used in cyber security

- Deep learning is a machine learning approach that integrates natural systems in various levels to incrementally educate from information.

- The improvements in the deep learning technologies has increased the possibilities of utilizing machine learning approaches to address the problems in various domains.

- Deep learning has been applied towards number of use cases related to cyber security like identifying, malware detection, malware classification, android malware detection, phishing and spam detection.

- List of deep learning methods used in cyber security:
  - Deep belief networks.
  - Convolution neural networks (CNN).
  - Restricted Boltzmann machine.
  - Recurrent neural networks.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Deep learning v/s restricted Boltzmann machines

Deep Auto encoder	Restricted Boltzmann Machines
It is a type of unmonitored neural network that takes a vector reference and tries to fit the response to the same variable.	It is a two-layer, bipartite, undirected graphical models that from the building blocks of DBNs.
These are flexible due to their controlled learning of condensed data encryption.	RBM's are unsupervised as Deep auto encoders and can be trained one layer at a moment.
These will reduce the computational resources to build an effective model by training one layer at a time.	In this type of network there is no interlayer connections but every nodes are fully connected as shown in Figure 4.
When the hidden layer has lower dimensionality than the feedback and production layer, the network is used to encrypt the information.	Restricted Boltzmann machines are deterministic, i.e. rather than definite values they give possibilities.
To gradually compress the information multi layers of auto encoders can be trained in series this is called stacked auto encoder	The design is conditioned by taking and supplying binary source information through the cycle. Then, the reconstruction of the source information is fed back through the design. The program's power will then be measured also utilized to adjust the scales. This method continues until the convergence of the system.
Sparse auto encoder comprises in the feedback and production surface of more secret points than there where only the secret level part is enabled at a given period.	Deep neural network can be created by stacking auto-encoders and RBMs, these are referred as stacked RBMs.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Convolution Neural Networks (CNNs)

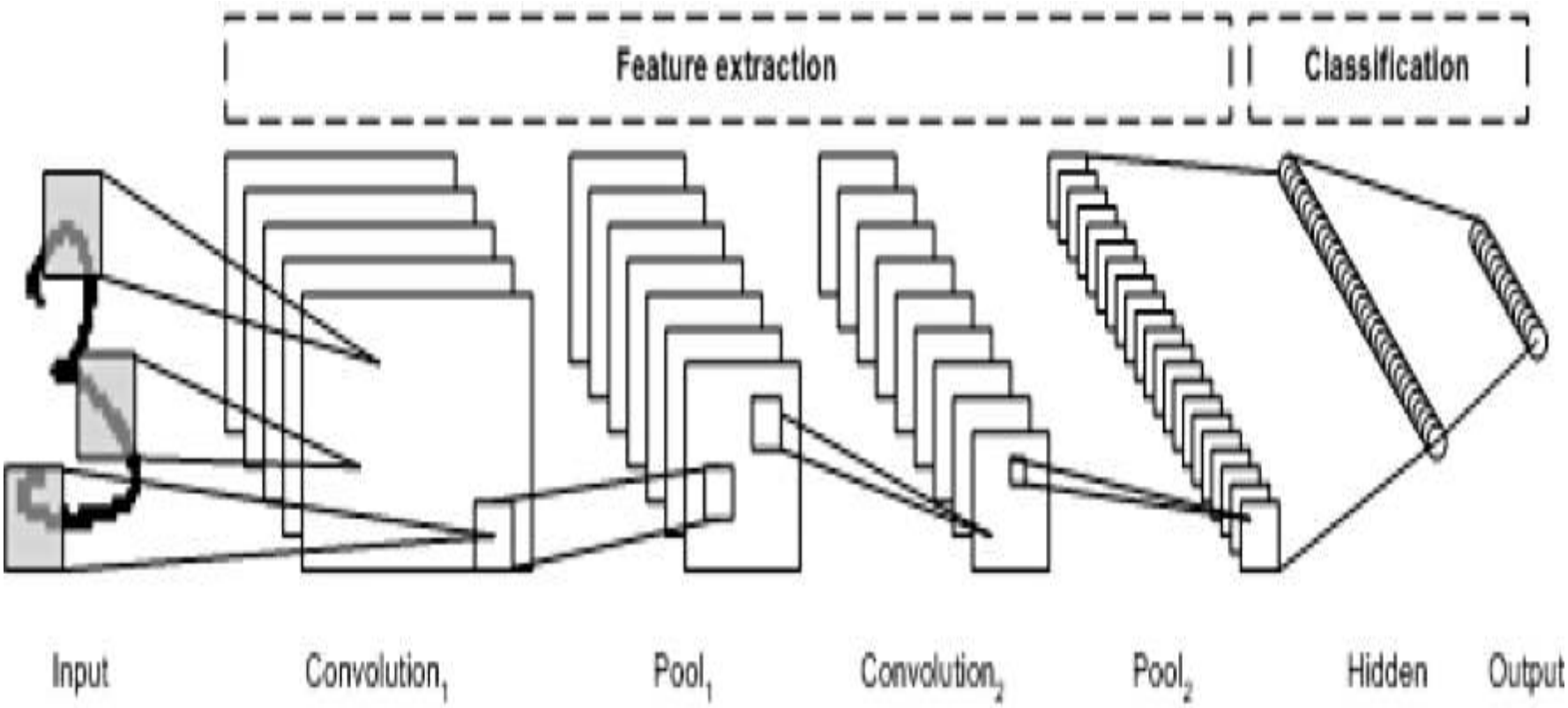


Figure: Convolution neural networks

Source: Smart insights © March 2018 the financial brand

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Recurrent neural networks

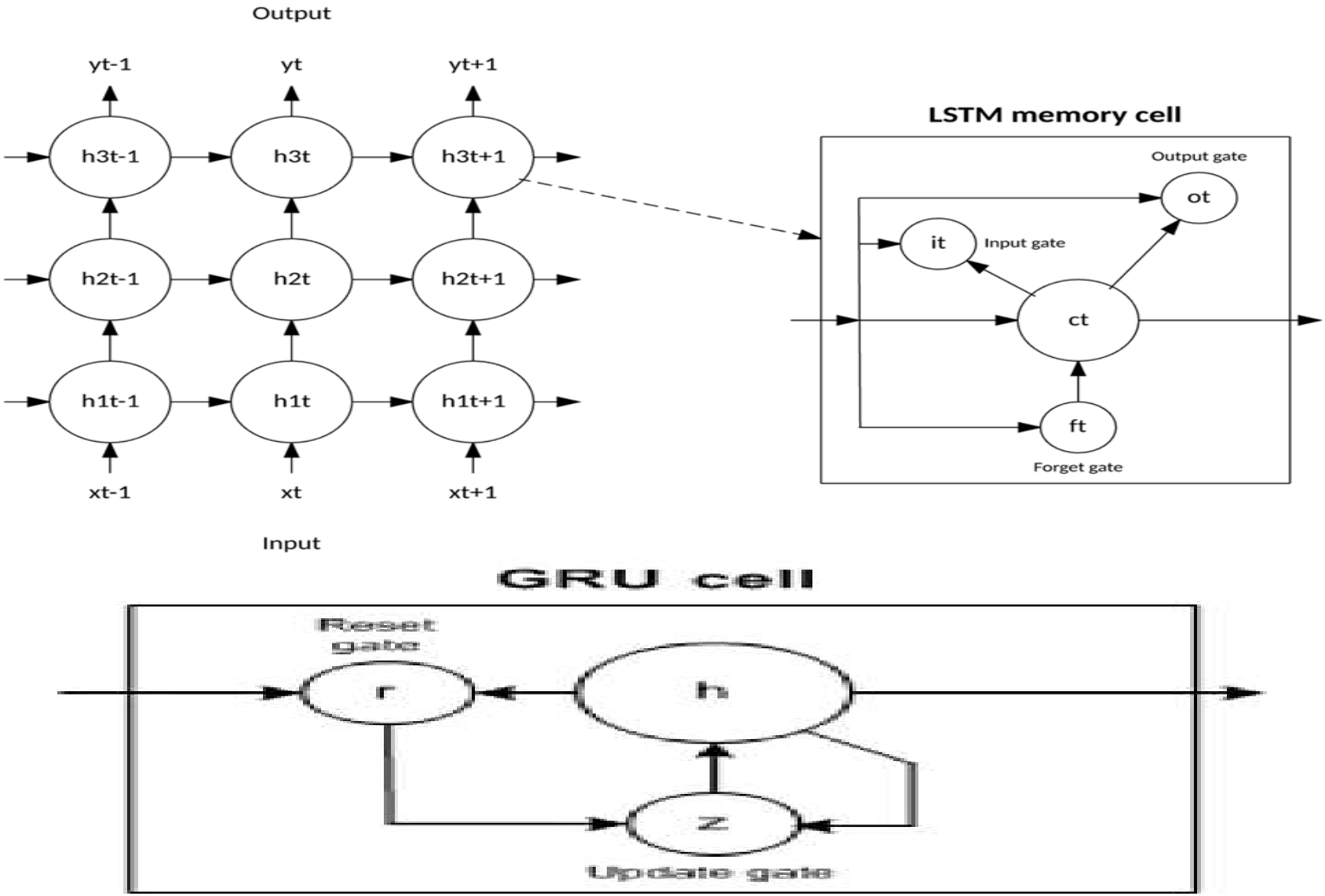


Figure: Recurrent neural networks

Source: Smart insights © March 2018 the financial brand

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Machine learning techniques: Loan underwriting & sentiment/news analysis

IBM ICE (Innovation Centre for Education)

Challenges	Potential Solutions
Verification of discrepancies in the document **Variety of document formats to checked for its accuracy. If performed manually, increase the likelihood of errors and mistakes. Also, it is time consuming**	ML based Natural language processing, along with image analysis leads to effective digitization of the entire process reducing potential errors and reducing the time.
Issues in the credit analysis: Documenting and analyzing past financial transactions is a tedious task if performed manually.	AI based metrics for credit analysis may lead to unbiased and validated credit-worthiness. This involves deployment of appropriate ML algorithm to process the credit history of the customer along with data analytics.
Assessment of debt-to-income ratio: Manual assessment of borrower's overall debt to income along with evaluation of their ability to repay the loan is cumbersome task.	ML enabled applications to evaluate debt-to-income ratio can be authentic source of evidence to make appropriate decision.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Sentiment or news analysis

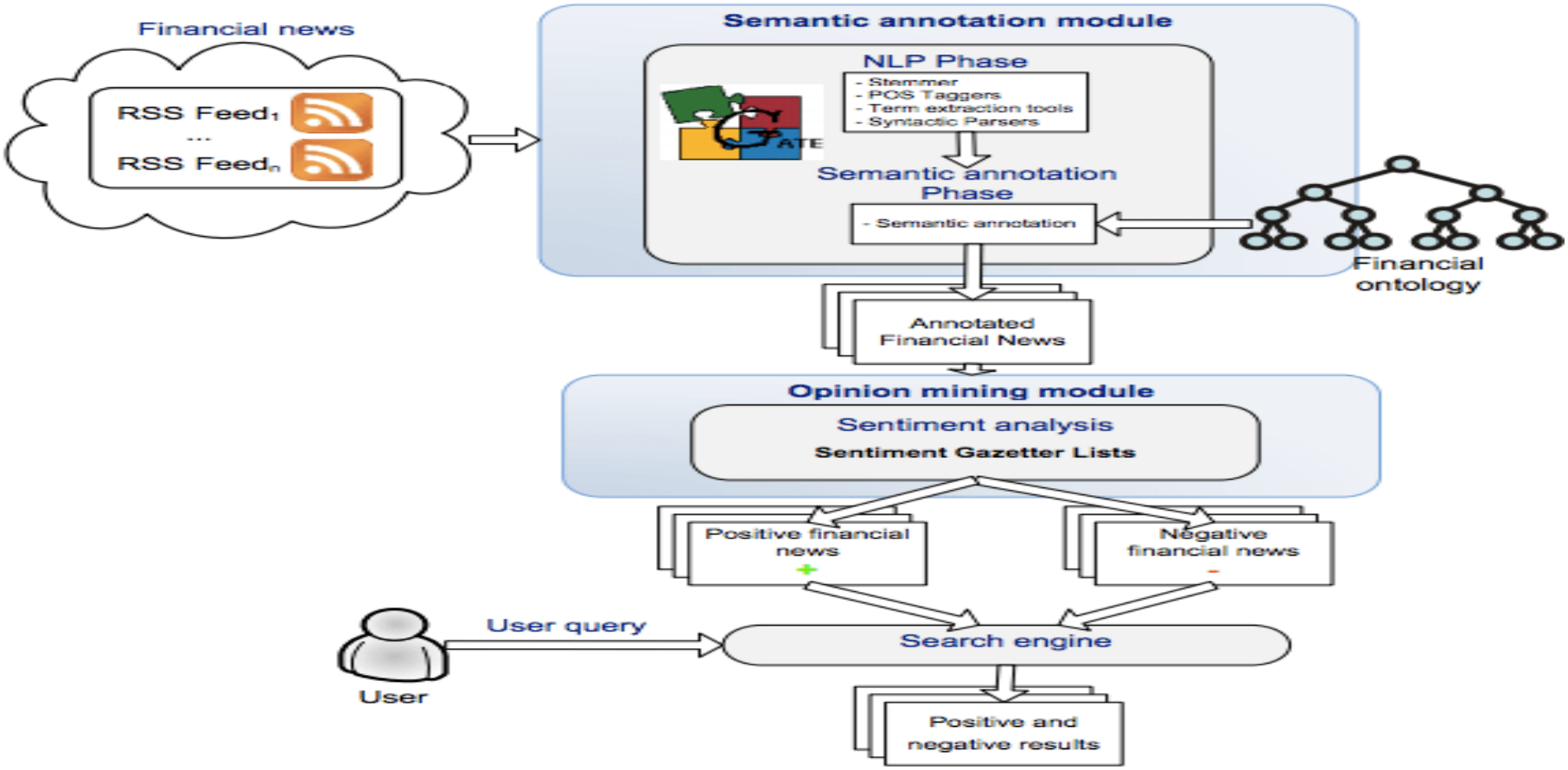


Figure: Flow of sentiment analysis application in the financial world.

Source: Semantic-Based Sentiment analysis in financial news, Vol-862

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Current challenges and opportunities: Banking and security domain

- Challenge: Lack of skills and data.
- Opportunity: More usable AI is coming.
- Challenge: Adoption.
- Opportunity: Professionals recognize the wider value.
- Opportunity: AI can simplify transparency and explain ability.

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Checkpoint (1 of 2)

Multiple choice questions:

1. Which of the following model is supervised learning model?
   a) Clustering
   b) Packing
   c) Regression
   d) Dimensionality reduction

2. Which of the following is a mathematical system which can be indicated by specific acyclic chart, it is also referred as a belief network?
   a) Deep neural network
   b) Bayesian classifier
   c) Naïve Bayes classifier
   d) Decision tree

3. Which of the following type of cyber attacks inundate systems with traffic to consume resources and bandwidth and make them unable to perform?
   a) Denial Of Services (dos)
   b) Phishing
   c) Watering hole
   d) None of the above

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Checkpoint solutions (2 of 2)

Multiple choice questions:

1. Which of the following model is supervised learning model?
   a) Clustering
   b) Packing
   c) Regression
   d) Dimensionality reduction

2. Which of the following is a mathematical system which can be indicated by specific acyclic chart, it is also referred as a belief network?
   a) Deep neural network
   b) Bayesian classifier
   c) Naive Bayes classifier
   d) Decision tree

3. Which of the following type of cyber attacks inundate systems with traffic to consume resources and bandwidth and make them unable to perform?
   a) Denial Of Services (dos)
   b) Phishing
   c) Watering hole
   d) None of the above

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Checkpoint (1 of 2)

Fill in the blanks:

1. _____ is a set of activities undertaken to prevent money or property from being obtained through false pretences.
2. _____ learning algorithms are task driven and the learning is based on a general rule of data classification.
3. Euclidean distance can be calculated as ----------.
4. ---------(also known as opinion mining or emotion AI) is a sub-field of NLP that tries to identify and extract opinions within a given text across blogs, reviews, social media, forums, news etc.,

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Checkpoint solutions (2 of 2)

Fill in the blanks:

1. Fraud detection is a set of activities undertaken to prevent money or property from being obtained through false pretences.
2. Supervised learning algorithms are task driven and the learning is based on a general rule of data classification.
3. Euclidean distance can be calculated as **$(x, x_i) = \sqrt{(\text{sum } (x_j - x_{ij})^2))}$.**
4. **Sentimental analysis** (also known as opinion mining or emotion AI) is a sub-field of NLP that tries to identify and extract opinions within a given text across blogs, reviews, social media, forums, news etc.,

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Question bank

Two mark questions:

1. List the challenges in banking sector and securities.
2. What is fraud detection in banking sector?
3. What is risk modelling in banking sector?
4. List the widely used machine learning algorithms in banking sector.

Four mark questions:

1. Describe applications of machine learning in customer data management in banking sector.
2. Explain with an example the applications of machine learning algorithms in personalized marketing.
3. What is risk management in banking sector and how it can be addressed?
4. Explain working principal of sentiment analysis (API).

Eight mark questions:

1. Explain with an example the efficient fraud detection in banking sector.
2. Write a short note on role of machine learning in addressing the challenges of banking sector and securities.

This document is available free of charge on studocu

Downloaded by Rohit Srivastava (rts080185@gmail.com)

Unit summary

Having completed this unit, you should be able to:

- Learn about machine learning in banking sectors, challenges in banking sector and fraud detection system in banking sector
- Understand manage customer data and algorithms in banking and security
- Gain knowledge on deep learning technology for personalized marketing
- Learn about applications of machine learning classifiers in credit risk analysis
- Understand the importance of machine learning in portfolio management systems
- Learn about algorithmic trading and stages for implementing applications of ai in marketing
- Deep learning approach for sentiment analysis of customers and customer services
- Understand cyber security in banking sector and loan underwriting and sentiment/news analysis
- Understand current challenges and opportunities in implementing machine learning technologies in banking and security domain

Downloaded by Rohit Srivastava (rts080185@gmail.com)