

## CHAPTER 13

# Overview of Machine Learning in Finance

A few years back I was approached by the financial client from the Southeast Asia region to help them with their machine learning effort since they were newly implementing it in their industry and they had become stuck with the practical implementation of the machine learning algorithm in their financial advisory services domain. During the implementation, I studied the financial industry around the world in order to get a better grip on what was required in order to implement this assignment.

Even today the implementation of machine learning and artificial intelligence in the field of the financial industry holds a lot of scopes.

In a very broad sense, financial services [1; [https://en.m.wikipedia.org/wiki/Financial\\_services](https://en.m.wikipedia.org/wiki/Financial_services)], as defined by Wikipedia, is comprised of businesses that manage money. These include credit unions, banks, credit card companies, insurance companies, accountancy companies, consumer finance companies, stock brokerages, investment funds, and individual fund managers. Of course, in our discussion, I am not going to talk about the government-related financial services.

As far as the segmentation of the financial services market is concerned, I will be using the following broad segments that are described in Table 13-1.

The clear differentiation in the business approach can be seen from Table 13-1.

**Table 13-1.** *Different Segments of Financial Industry, Broad Classification*

Type	Focus	Services and Differentiation
Consumer Finance	Providing personal loans, indirect financing, including lease and sales financing, credit card issuers, pawn shops, and payday loan providers	Provide unsecured loans to individuals for both commercial and personal purposes. They operate in the subprime category and provide loans to individuals with bad credit history
Capital Markets	Undertake activities, including trading, brokerage, strategic advisory, portfolio management, asset management, and investment advice	Primarily work as intermediaries, either to provide or manage capital, thereby satisfying financial goals of institutions and individuals
Diversified Financial Services	Providing financial services, such as financial advisory firms, investment research companies, stock exchanges, and credit rating companies	A majority of their revenue is derived from one particular area of expertise and specialized activity

In Table 13-1 the market segmentation of the financial services industry [2; <http://media.lockelord.com/files/upload/advisenIndustryReportDivFinConsFinCapMrkt.pdf>], which is based upon the financial services industry, subprime and credit crisis weighs heavy, according to a 2008 edition report by John W. Molka III, CFA Senior Industry Analyst and Editor. We first look at the consumer finance section, which is a group of companies providing personal loans, indirect credit card issuers, pawn shops, and payday loan providers. This list does not include mortgage and lending companies. The main item that these companies deal with is unsecured loans to individuals. Of course, all the statistics given in the report are based on the US; however, the percentages are applicable for all the developed countries. For example, the consumer finance companies hold 23 percent of the consumer credit as of 2007, and this can be applied to other developed economies as well. While this statistic looks dated by more than 10 years, it gives us an idea of how much share the consumer finance companies hold in the consumer credit in a nation’s economy.

Next in the segment are capital markets, which are comprised of organizations undertaking stock trading brokerages, strategic financial advisory portfolio management, asset management, and professional investment advice. These companies are intermediaries, and they work for the goals of other financial institutions or individuals.

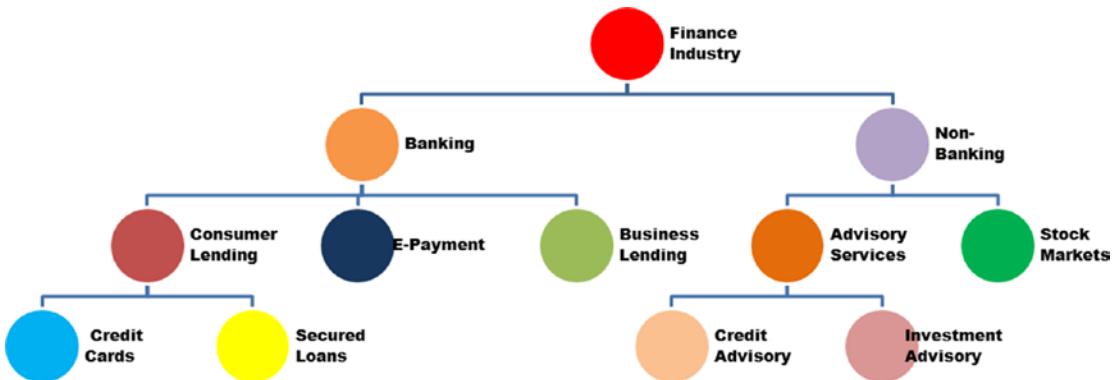
These companies include investment banking and trading in stocks, bonds, derivatives, and commodities. It also includes asset management firms that manage large funds for individuals and institutions around the world. The increasing trend in the sector is the rise in investment banking and securities dealings and that of securities brokerages. The largest services category is investment banking and security dealing, which comprises 41 percent of the total segment business. An increasing trend in the segment is the rise of mutual funds and investments, and this trend can be seen both in the developed and emerging markets.

Which is essentially the companies involved in financial advisory investment research, stock exchanges, rating, and credit rating. This is a market where there are top 10 players, like Citigroup, Berkshire Hathaway, Goldman Sachs, Morgan Stanley, JP Morgan Chase, and Merrill Lynch. The next is the diversified financial services segment.

Now that we have looked at the broad segmentation in the financial industry [2], we will now look at some of the key segments and the value proposition that they provide.

## Financial Segments

In the drill-down financial statements diagram in Figure 13-1, we can see the categorization of financial industry into banking and nonbanking segments.



*Figure 13-1. Financial segments*

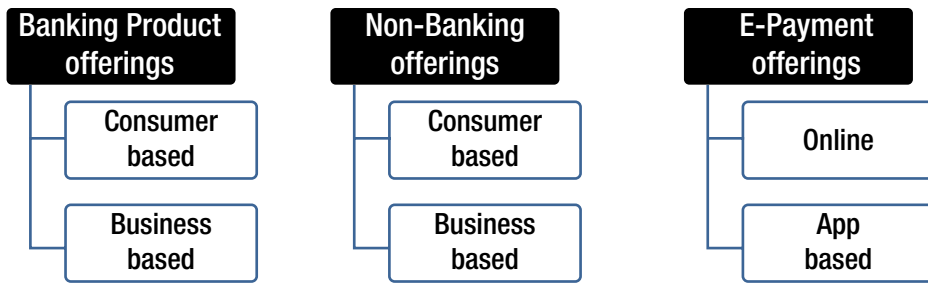
The banking segment is comprised of consumer lending, which is a very structured form of business where the creditworthiness of the consumer is ascertained through the creation of credit profiles of the consumer. If the credit profile is found to be risky, then the rescue level is ascertained, and only then is the lending possible in the

segment. Next is the business lending, which is also a very structured form of lending to commercial organizations for their investment and growth. Although this type of business lending also has a lot of variations, I am not going to go into them now. The third type of banking segmentation is that of e-payments. Both the consumer and the business make electronic payments through the electronic payment banking systems, which are local to each country. The consumer lending can further be divided into credit card lending and secured loans lending. Credit cards are issued by banks only when the creditworthiness of a person is ascertained so that they can perpetually use the credit and then repaid within the given stipulated period. Secured loans are given only when the creditworthiness of an individual is ascertained with respect to the assets that they hold and the liabilities that they have against various loans that they might have taken in the past.

In the non-banking space, we divide the segment into advisory services and the stock markets. In advisory services we again divide our segment into credit advisory services and investment advisory services. The credit advisory services are related to the credit ratings of agencies, banks, countries, and companies. There are various credit rating agencies that operate all over the world and they fall in this segment category. Investment advisory is concerned with management and advice on the funds held by various individuals and organizations. This includes portfolio management services and other such financial advice given by various agencies. Stock markets involve a segment of portfolio management services and brokerages that can be the regulated stock exchanges and unregulated stock exchanges, like the ones used by blockchain technology (e.g., bitcoin). We will be referencing this drill-down version of our segmentation of the finance industry in this finance section throughout this book.

## Finance Value Proposition

Value proposition, as defined by Google dictionary, is marketing and innovation, service, and features intended to make a company or product attractive to customers.



**Figure 13-2.** *Finance customer value proposition*

The finance-based services that are offered by organizations have a different values proposition for consumers and businesses for both banking and financial services. This is depicted in Figure 13-2.

In a typical financial organizational setup, there are some common departments like consumer services and business and investment services. Especially in the consumer services division, it is further divided into general consumers and premium customers or privileged customers. The financial institution like a bank or a financial advisory or portfolio management company differentiates between the consumers in this way and offers its different level of services accordingly. We will now look at how machine learning is being applied in some of these departments. An important thing to note is that the customer generally faces only the marketing department or the customer support department, and the rest of the department works in the back end. However, in the financial services industry, there is an increasing trend to have a dedicated relationship manager for privileged or premium services customers.

In Figure 13-1 we already saw the type of services under non-banking segment, mainly comprised of advisory services (both credit and investment) and the services related to the capital of the stock markets. It is important to understand here that both credit and investment advisory services can be for personal individuals as well as for other companies, organizations, or institutions. The consumer-based are the advisory services provided to individuals as customized or tailored for the needs of that individual. These types of credit advisory services are premium services meant for the privileged people of the society. The focus is to provide credit information such as rating, bankruptcy status, financial analysis, fraud alert, and other services in this category. The investment advisory services are provided in the form of portfolio management, asset management services, and fund management services for individuals. It is upon the individual to buy both advisory and management services from that particular provider or to buy advisory services from one provider and the investment management service

from another provider. Usually, consumers or individuals who have high net-worth generally prefer to get investment advisory and management services from the same financial service provider.

There is another segment or category of financial services that have value proposition falling under the category of e-payment. These are services provided online through the internet on various websites and also in the form of app-based services through mobile phones such as Android or iPhone. The offerings in this category of e-payment or online payment financial service providers [3; [https://en.wikipedia.org/wiki/List\\_of\\_online\\_payment\\_service\\_providers](https://en.wikipedia.org/wiki/List_of_online_payment_service_providers)] is more oriented toward country-based providers, and there are very few providers who can actually call themselves global. Prominent in this category are providers like Paypal, Payoneer, Apple Pay, Amazon Pay, and Google Pay, who are from the US but have a major global presence. PayU is another that has a good global presence; however, we have some regional ones, such as Trust in the European Union and Alipay in China. This list is not exhaustive; however, there are many more that have deep regional and local presence.

The reason for such a thin growth in the global players of this segment is the fact that financial services are a highly regulated industry. The major central banks around the world regulate and bring out norms, guidelines, and proposed laws for regulating these types of services. Any major operator who wants to become a global operator has to have not just the understanding but the capability to implement regulations in major countries around the world where it wants to operate. The global players that I have listed have been able to meet these norms, and hence they have been able to build a global base.

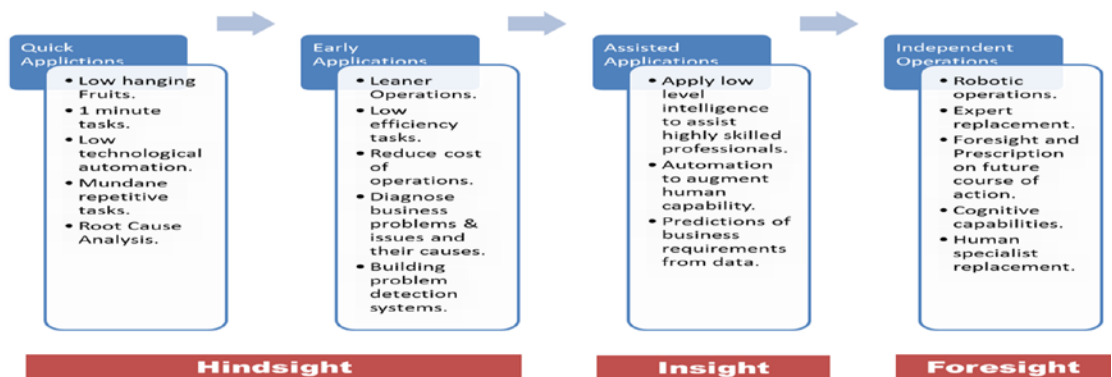
There are some service providers that have only an online presence; however, there are some that have an app-based mobile presence as well, such as Alipay from China, which has both mobile and online payment platforms. Payoneer provides e-wallet services or digital payments services, which allow customers to send and receive funds into their bank accounts or reload them into prepaid debit cards. This company focuses more on business-to-business payments across the world. It is now known as the backbone of the business-to-business digital transaction system globally. Similar to Payoneer is PayPal, which allows online money transfers as a replacement to checks and money orders. Paypal, however, does not focus just on business-to-business but also on business-to-consumer transactions as well.

Before I wrap up the value proposition section, I would like to mention the recent developments that have happened in the e-payments category, which have been the

advent of blockchain technology that have brought virtual currency like bitcoin and ethereum. These currencies have not been accepted by major countries and their central banks as a mode of payment and have been banned in some countries like South Korea and India. I don't want to go into the controversy of these e-payment services; however, you should be aware that there are a lot of startups around the world that have cropped up and are providing digital payment services in these blockchain currencies as well. Prominent among them is Payza, which allows bitcoin money orders and transfers to credit or debit cards via bitcoin or altcoin. In my opinion, this perception that a lot of illegal transactions happen through the blockchain type of currencies will get moderated in the near future and will come under the regulations in major countries around the world. Blockchain technology that is peer-to-peer or consumer-to-consumer is more trusted by the users since there is no centralized server or centralized institution that is controlling the entire system. The decentralized financial framework, in my opinion, is the future that is waiting to happen once the technological understanding deepens with the regulators around the world and people adopt the technology in spite of the regulations and the warnings given by various governments.

## The Process of Technology Adoption in the Finance Sector

Figure 13-3 depicts the process of technology adoption taken from this book's healthcare section.



**Figure 13-3.** *Process of technology adoption*

**Descriptive Analytics:** This field of analytics is invoked to know the answers to questions for projects that have already happened in the past, such as what is the status of X financial transaction in the past?

Traditionally the finance sector has been using descriptive analytics to produce reports about its customers for any given services category, and this has had a big impact with the top management while making various decisions, such as whom to target for any new financial product, how do we make strategy change decisions against competitors, and how is our service perceived by the end customers? All these questions are very well-answered by descriptive analytics. This is beside the trend of putting together the five-number summary for any data to the management [4; [https://en.wikipedia.org/wiki/Five-number\\_summary](https://en.wikipedia.org/wiki/Five-number_summary)]. For example, for comparison of financial data sets between the spending patterns of the millennials and the baby boomers generations, five-number summary may tell us about how the minimum spending pattern and maximum spending pattern differ between these generations. It then becomes easy for the top management to customize their financial services offerings to them specifically. Unless descriptive analytics is used, it will not be possible to take out such information.

Product-wise sales reports, product-wise discount reports, and store-wise and product-wise sales reports are some of the examples where descriptive analytics has been used traditionally. This trend is not going to go away. We will see the use of descriptive analytics in the retail sector. However, we will also see the emergence of other analytics that I talk about in the next section. I give you an example of how to use the five-point summary of financial data in the short example Python code in Listing 13-1. Since this is an example, I will be taking not more than 10 rows of data; however, in the real world this would be comprised of data for millions of users. All values for spending are in US \$ and yearly value.

**Listing 13-1.** Python Code for Five-Points Summary

```
import numpy as np
bboomers = np.array([14230, 345, 1912, 472, 63, 861, 270, 713])
fivepoints= [np.min(bboomers), np.percentile(bboomers, 25,
interpolation='midpoint'), np.median(bboomers),np.percentile(bboomers, 75,
interpolation='midpoint'),np.max(bboomers)]
for fivepointsummary in fivepoints:
print(fivepointsummary)
```



Output:

```
63
307.5
592.5
1386.5
14230
```

```
import numpy as np
millennials = np.array([12519, 845, 912, 72, 93, 615, 70, 538])
fivepoints= [np.min(millennials ), np.percentile(millennials ,
25, interpolation='midpoint'), np.median(millennials ),np.
percentile(millennials , 75, interpolation='midpoint'),np.max(millennials)]
for fivepointsummary in fivepoints:
    print(fivepointsummary)
```

Output:

```
70
82.5
576.5
878.5
12519
```

After running the Python code in Listing 13-1, we see the output that the minimum average yearly spending by baby boomers in our data set is 63, that the 25th percentile happens at 307.5, the median lies at 592.5, the 75th percentile is at 1386.5, and the maximum that a baby Boomer generation spends in our data set is \$14,230 yearly. In the second code snippet for Listing 13-1, we see the minimum that millennials spend is \$70, and the 25 percentile stands at 82.5, the median is at 576.5, the 75th percentile is at 878, and the maximum spend is at \$12,519 yearly. In a simple comparison like this, an excellent guide to determine who is spending more for example, in our data set the baby boomers generation is spending the highest; however, they are also spending the lowest, so the dispersion is higher in the case of baby boomers than the millennials. Baby boomers spending appears in both max and min function in Listing 13-1. You will find this kind of a scenario in practical datasets. The Key thing is to know how to interpret and analyze it. The median values speak a lot about how the spending of the baby boomers and millennials in our data set is structured, and the median value for the baby boomers is pretty low, at 82.5, in comparison to the baby boomers, which is at 307.5. This means that the baby boomers are indeed higher spenders, with the midpoint line much higher in our data set.

**Diagnostic Analytics:** We know that this type of analytics is used to know the root cause of a phenomenon, such as a project success or a failure. Why did the X fraud happen? What was the root cause of this fraud to happen in our financial systems? What was the root cause that our financial systems did not catch the money laundering transactions? All such questions can be answered using diagnostic analytics. All forms of financial risk assessment come under diagnostic analytics. A practical implementation is a tool by Pricewaterhouse Cooper known as Halo [5; <http://halo.pwc.com/>]. This tool works on accounting journals and ledgers and has the capability to flag higher-risk transactions. Of course, this is not an online real-time solution since the effort is only to diagnose and point out the financial risks. Accounting is about old transaction records, and that is what a tool like a halo tries to extract information from a client's systems and discover to bring forth high-risk transactions with their patterns and trends. It also gives out information on process inefficiencies in the accounting process, like eliminating duplicate accounting journal entries, etc. This is especially useful for a large financial organization because of the high volume of transactions that happen during a particular period in various departments.

**Predictive Analytics:** We have seen in the healthcare and retail sections of this book that this type of analytics is used for determining the outcome of an event in the future, such as a financial transactions success or failure, financial budget overrun, or a budget slippage for an organization such as a bank or a company. This analytics is something that brings about the actual ability to avoid situations in the future—for example, by using predictive analytics we can predict if a particular type of financial transaction is going to succeed or fail. There could be various parameters that the machine learning model code takes into account for analysis before coming to such a conclusion. However, the whole purpose is to give a sense to the top management on the direction that the financial side of the business is going to take. Even this does not solve the problem, but it merely tells or acts

as a warning signal of a failure or a deviation from the plan in the future related to finances. When we apply predictive analytics, then we get a lot more fields of finance that come into play, such as cost accountancy, managerial economics, and decision support systems. Giving out predictive costing reports based on past cost accountancy data is something that I have implemented for one of my clients. This helps large financial businesses stay on track with their cost budgets and ensure the right flags are raised in case the actual cost is predicted to increase as a result of prediction by a machine learning model. In management accounting or economics, as it is known in the academic circles, the focus is on what-if scenarios, such as \$x million invested in pqr and abc projects and which one is better. And why? What if we invested the \$x million in acquiring a certain startup? What would be the ROI on that investment in comparison to putting in the above two projects? All such complex decision-making questions are very easily answered by building machine learning models based on past data and applying them based on principles of management accounting.

**Prescriptive Analytics:** We now know that in this field of analytics the maximum value of analytics is achieved as it builds upon the prediction made based on predictive analytics, and it prescribes actions that should be taken for the future. In prescriptive analytics, it uses machine learning to build a future plan of action that is based on past failures in financial transactions in similar situations, such as failed financial transactions, and it gives out specific activities that need to be carried out in order for the financial transaction to be a success. In the financial sector, prescriptive analytics is now truly able to do course corrections even before the failures happen. Every financial organization around the world is striving to make a breakthrough in this field of analytics; however, there are very few successes as of the time of writing this book. I say this because if you take the case of business cycles, although it is known from the domain of macro-economics that they occur roughly 8 to 10 years, nobody is still able to

predict when they would really happen again. The prescriptive part starts when there is an accurate prediction made; then the actions that need to be taken even before this happens come after that. For example, a global financial conglomerate bank creates a machine learning model to predict the onset of the next global financial recession with an accuracy of 92 percent; however, this is not enough. The top management needs to know what steps they need to take so that the impact on the organization is nearly minimal. Building and creating such a system would need a huge amount of data and many years of a model building-testing-feedback cycle in order to arrive at production. There are organizations that are trying to build models to predict the crash and boom cycles in capital markets; however, due to the involvement of such a high amount of data and environment variables, there is no breakthrough yet.

## End Notes

- [1] Financial Services. [https://en.m.wikipedia.org/wiki/Financial\\_services](https://en.m.wikipedia.org/wiki/Financial_services)
- [2] John W. Molka III, CFA Senior Industry Analyst and Editor 212-984-2753, FINANCIAL SERVICES INDUSTRY SUBPRIME AND CREDIT CRISIS WEIGHS HEAVY, <http://media.lockelord.com/files/upload/advisenIndustryReportDivFinConsFinCapMrkt.pdf>
- [3] List of Online Payment Service Providers, [https://en.wikipedia.org/wiki/List\\_of\\_online\\_payment\\_service\\_providers](https://en.wikipedia.org/wiki/List_of_online_payment_service_providers)
- [4] The Five number summary, [https://en.wikipedia.org/wiki/Five-number\\_summary](https://en.wikipedia.org/wiki/Five-number_summary)
- [5] Halo software by Price waterhouse Coopers, <http://halo.pwc.com/>