DIMITAR A. TSENOV ACADEMY OF ECONOMICS - SVISHTOV

DEPARTMENT OF MARKETING

Dimitar Georgiev Trichkov

PRODUCT DIVERSIFICATION USING BIG DATA ANALYTICS

AUTHOR'S ABSTRACT

Of a doctoral dissertation

For the award of the doctor's degree in Economics (Marketing)
Field of higher education 3 "Social, economic and legal sciences",
professional field 3.8. Economics,
doctoral programme "Marketing"

Scientific Advisor:

Assoc. Prof. Todor B. Krastevich, PhD

SVISHTOV

2021

The dissertation has a total of 218 pages and is structured as follows: Introduction (6 p.), three Chapters (202 p.), Conclusion (4 p.), and a List of References (6 p.). The thesis contains 46 tables and 50 figures. The bibliography includes 107 sources - in Bulgarian, English, and Russian.

The dissertation was discussed and referred for defense pursuant to procedure under the Act on the Development of the Academic Staff, as implemented by the Department of Marketing at the D. A. Tsenov Academy of Economics at a meeting held on December 14, 2021.

The open meeting of the scientific panel for the defense of the dissertation is scheduled at 11.00 a.m. on February 25, 2022, in the Rectorate Conference Hall of the Dimitar A. Tsenov Academy of Economics – Svishtov.

The relevant materials for the defense of the dissertation are available to the stakeholders at the Department of Doctoral Studies and Academic Development of the Dimitar A. Tsenov Academy of Economics –Svishtov.

Table of Content

I. General characteristics of dissertation

The relevance of the research topic

In recent years, technological developments have brought closer together previously unrelated areas of economic activity, such as telecommunication and financial services. Mobile devices are becoming the main means of communication and performing an increasing range of financial and other services, and the mobile traffic generated by this consumption is stored in each individualized specific customer profile in the databases of telecommunication service providers. The existence of such arrays and the supposed internal dependencies suggest their exploration and subsequent generation of innovative ideas for new business models. The challenge for their development is the internal conservative corporate nature of telecom and traditional banking institutions, where there is a traditionally established corporate culture that hinders agile and operational asset and process management, following the example of start-up companies, for the implementation of alternative technologically anticipatory services. Illustrative of this assertion is the fact that so-called fintech companies are increasingly successful in competing with traditional financial institutions in providing financial services, and telecommunications giants are no longer perceived as innovation leaders, but rather as utility companies that, similar to those in the electricity and water sectors, provide communication and connectivity services. Such a development is unacceptable for both industries, which would understandably have a financial and image interest in being leaders in economic activity associated with high quality, technology and stability. One possibility for this is through the creation of a specialized unit (in-house or legally independent), telecom and banking institution to jointly utilize their existing assets, delegate rights and obligations to this structure, which with the necessary flexibility, independence, competence and speed to successfully compete in the innovative services sector with digital start-ups, exploiting the advantages created by the synergistic effect of the combined power and scale of its related parent companies.

Object and scope of the study

The **object** of this study are the alternative services, out of the core telecommunication business, offered by telecom companies. The **scope** of the study is the possibility of market growth and the associated additional financial flows generated as a result of the telecommunications company offering its own digital financial services in partnership with a financial institution.

The thesis supported by the author is that telecommunication companies have a huge potential to diversify their service offerings based on the big data generated by and in relation to the activity of their customers.

Aim and objectives of the dissertation

The aim of the research is to analyze the possibility of increasing the revenues of the telecommunication company through diversification, consisting in offering own financial services, outside the usual telecommunication activity on the basis of the available technological, information and physical infrastructure of the enterprise.

To achieve the stated objective, we set the following research objectives:

- Demonstrate the ability of telecommunications companies to add value to financial sector service offerings. The vast amount of behavioral data available in telecom repositories combined with traditional financial data can improve individual customer scores, thereby not only contributing to the development of innovative business models, but also serving society by providing financial services to customers for whom they have been inaccessible or limited throughout traditional channels.
- Demonstrate the benefits of cross-industry use of big data and sharing analytical results with different application areas. In their daily activities, people are using mobile communication devices for an increasing variety of purposes and services. This usage generates a huge volume of information that is stored and processed by their provider for their own purposes. Our task is to demonstrate the possibility that this data can also be used to create and manage innovative business models.
- To develop predictive models for analyzing behavioral telecom data and the possibility of using it in an unrelated economic domain.
- To apply customer segmentation based on behavioral data and test the hypothesis of its adequacy in the absence of other information. Individuals' access to borrowed funds is fundamental to the development of the economy and society. In many cases, especially for young people and in third world countries, historical financial information is lacking for some customer segments, and therefore a significant proportion of them are deprived of access to financial services. The method proposed here postulates that the availability of telecom consumption data can successfully replace the lack of traditional data.

- Demonstrate a business model involving the active participation of a telecommunications company in the financial services market. The development of digital and mobile technologies has made telecom companies a major platform for the development of various services, including financial services. In order for companies to fully exploit their vast potential and infrastructure, they need to dissociate themselves from the role of a utility company (similar to electricity and water providers), and take the risk of being innovators in areas of activity outside of their usual one.

Coverage and limitations adopted

The following limiting conditions were adopted in the present research:

- *Research and analysis of individual customer data only.* Customers of telecommunication companies are both individuals and business organizations (legal entities). In cases where a customer owns more than one subscription, the relationship between the owner and the service user is ambiguous and not guaranteed. Therefore, only customers with a single mobile service subscription are included in the study to guarantee correct data interpretations.
- *Product diversification is considered in the context of financial service provision.* Exploring the potential entry of telecommunication companies into innovative business models in different economic areas is an objective opportunity which scope could not be covered by this dissertation. Alternative business models in the financial sector are attractive due to their scale, the availability of direct communication with end-users by the telecom and technological developments allowing the combination and complementarity of services from the two sectors under consideration.
- *The use of big data analytics for the purpose of alternative business models is considered only in economic terms.* The telecommunications and banking sectors are industries subject to strict and stringent legal regulations. This study does not consider and account for the impact of regulations in terms of licensing regimes at work, as well as data protection restrictions (GDPR) in the EU and cyber security issues. These aspects of activity could be the subject of a separate study.

II. Structure and content of the dissertation

The dissertation consists of an introduction, three chapters and a conclusion with a total length of 221 pages. It is structured as follows:

INTRODUCTION

CHAPTER I PERSPECTIVES AND CHALLENGES IN TELECOM AND FINANCE

1.1. CHALLENGES AND OPPORTUNITIES FOR TELECOM COMPANIES IN DIGITAL SERVICES

1.1.1 Overview of the current state of telecom operators in the context of the digital economy

1.1.2. Data organization in a telecommunication company

1.1.3. Telecommunication companies and mobile financial services

1.2. CHALLENGES IN THE DEVELOPMENT OF FINANCIAL INSTITUTIONS IN THE CONTEXT OF DIGITALIZATION

1.2.1. Banking institutions in a changing environment

1.2.2. Financial institutions- digital transformation and innovation

1.2.3. Non-bank financial institutions, fintech and P2P deposit and funding platforms

1.3. PRODUCT DIVERSIFICATION AS AN OPPORTUNITY FOR CUSTOMER CAPITAL GROWTH

1.3.1. Change management in development strategy

1.3.2. Resource-based approach as a policy to improve performance

1.3.3. Diversification as a means of increasing customer capital

1.3.4. Telecom and banking convergence

CHAPTER II FINANCIAL SERVICE PROVISION AND CREDIT RISK ASSESSMENT

2.1. CREDIT RISK ASSESSMENT

2.1.1. Credit information reporting systems

3.3.2. Added value of telecom - banking partnership

3.3.3. Business model for joint implementation of mobile financial services

CONCLUSION

Sources Cited

Introduction

In the introduction of the dissertation, the relevance and importance of finding new application models and alternative uses for the vast amount of information entering and being processed in the databases of communication service providers are briefly outlined. The object, subject, aims and objectives of the research and its limiting conditions are set out. The methodology and tools used in the study are outlined.

Chapter One: Prospects and Challenges in Telecommunications and Finance

Chapter One provides an overview of the state of the telecom and banking industries and the challenges associated with the digitalization of the economy. Market threats, including those from alternative providers, are examined. Possible shared development strategies are reviewed, as well as some case studies of successful cross-industry collaborations.

1. Telecommunication companies in terms of digital services

1.1. Ever since their emergence and the subsequent mass take-up of mobile communications, telecom operators have always been perceived as agents of progress, new digital technologies and innovation. For a long time, they were at the heart of digital transformation, first in voice telecommunications, but more recently also taking the next step towards offering comprehensive ICT services to both businesses and ordinary consumers. However, when comparing the Total Returns to Shareholders (TRS) and EV (enterprise value)/ EBITDA for the period 2005 - 2018, the telecom industry proves to be a less attractive sector for investors, which is far from the leading industries (Rajesh Duneja, 2020) on the above-mentioned indicators. This is mainly due to the sector's relatively volatile revenue levels over the last 10 years, combined with constant and often increasing capital expenditure based on constant technological development. As a result, access to borrowed capital on good terms, which the sector continuously needs to cover its technological

growth needs, is limited. To overcome these shortcomings, telecommunications operators need to analyze and realize the possibility of developing alternative areas related to their core business by reviewing the management and rational use of their available assets. For the moment, operators are most actively engaged in activities in the areas of IoT, M2M, smart homes, cities, mobile money and wallets, etc. For companies in the industry to achieve greater financial returns, they need to expand their services beyond offering connectivity and communications and move towards comprehensive offerings that support the digitization of the business to as a whole.

In order to adapt to an ever-changing environment and serious competition, the successful telecom of the future needs to focus its strategy on achieving both internal (organic) growth and external (inorganic) growth. For more than a decade now, companies have been using big data (Big Data) and big data analytics to enhance the quality of their products, customer segmentation, targeted campaigns, customer erosion predictions, last minute offers, etc. Due to constant improvements in technology, terminal devices used, how they are used, etc., the data sets available to operators are becoming significantly more voluminous, diverse and connected.

1.2. Data warehouse (DWH- Data Warehouse) is a major component of business intelligence (BI) systems of any organization, in which large databases of heterogeneous data with different sources are integrated, interconnected according to certain attributes. In a telecom DWH, millions of raw records are processed and warehoused per day, and the data analysis and processing is performed in real time using the Online Analytical Processing (OLAP) approach. Particularly common in the customer area is the use of customer data for customer segmentation, customer erosion analysis (so-called churn) and tailored individual optimization approaches, direct campaigns, connected selling offers and other customer relationship management (CRM) activities. The challenge for companies is the process of mapping different identities to one person: the integration of a single user profile: a link between customer profiles in different systems and environments, identifying the user as the same person. The unified profile allows companies to better understand their customers and interact with them, knowing who they are in different environments and what they are looking for in them. It makes it possible to analyze past behavior across different systems to build and personalize more effective customer interactions.

The application of big data analytics is not limited to pure marketing activities: when providing personal services, especially financial ones, it is particularly important to correctly identify the

user, as well as considering the likelihood of misuse or fraud. Undoubtedly, the vast amount of data generated by connected devices has its potential, and one of the challenges in using it is having data scientists who can skillfully extract the information needed for the business. Managing data quality is the other major challenge for vendors, requiring specialists to have broad business yet IT knowledge to construct the internal DWH architecture in an optimal way. A successful strategy in this direction would be to carve out an activity, both business and IT focused in nature, to perform the functions of big data collection, processing and analysis for needs other than business as usual.

1.3. In recent years, telecommunications companies have also been moving into the world of financial services. They are one of the fastest changing services worldwide, undergoing a significant transformation. The digitization of financial payments has reached levels where consumer preferences have clearly shifted from cash to the use of digital channels over the past five years, and in 2019, for the first time, digital transactions represent the majority of mobile money flows (Naghavi&others, 2019). At the end of 2020, mobile money offered as a service is available on the networks of 310 providers in 95 countries around the world (Andersson-Manjang & Naghavi, 2021)) is available in 96% of the countries in the world where less than one-third of the population has access to formal financial institutions. In the development of mobile finance, the role of telecom operators is still mainly to be the medium of use for their users. The measurement of financial success and profitability from mobile financial services for telecom operators cannot yet be compared to those generated by the core business (Vonthron & Almazan, 2014). Mobile finance is primarily a service impacting company OPEX and generating commissions for sales and marketing, the revenue from which cannot compete with the EBITDA margin of 25-35% typical of the core telecom business. Operators in Bulgaria, like their counterparts in Europe and around the world, offer mobile wallet services related exclusively to cashless financial transactions. The country's two leading telcos rely on their own digital wallets: pay by Vivacom[1] and A1 wallet[2], which complement rather than compete with premium banking services. Interest in such applications is growing, especially in view of the situation in 2020 related to the limited movement resulting from the COVID-19 contagion, with digital banks, e.g. Revolut

[1] https://www.vivacom.bg/bg/pay (5.11.2021 г.)
[2] https://www.a1.bg/a1-wallet (5.11.2021 г.)

with simple, clear and aggressive pricing, for the time being managing to attract more customers thanks to their distinct advantage.

2. Financial companies in terms of digital services

2.1. The global banking system is now, according to a Deloitte report (Val Srinivas, 2018), not only bigger and more profitable, but also more resilient compared to the period of the last ten years. Since the 1980s, commercial banks have been constantly engaged in technological innovations to make it easier for their customers to reach larger groups of people while optimizing their costs. In order to increase their profitability even further and at the same time free up financial resources to invest in the growing digitalization, financial institutions need to reduce their costs further. As the report (Jan Bellens, 2020) points out, in 2018 banks in Europe continued to reduce their physical presence on the continent as a consequence of the fact that a well-developed branch network is becoming less of a competitive advantage. The number of customers interacting with banks in a digital environment is growing steadily, at the expense of those being served in bank branches. The situation in Bulgaria follows global and European trends, and the BNB in its report (БНБ, 2019) highlights the conclusion of the internet's consolidation as a major tool influencing the optimization of costs of credit institutions and improving their overall efficiency. At the same time, it also points to the main risk of this development related to the increase in change-specific IT spending, such as protection against cyber risks of information systems digital finance and e-banking are taking on an increasing role and have the potential to displace traditional ways of interacting with customers even in a shortened timeframe. Even before the COVID-19 crisis, a PwC study (Bob Sullivan, 2014) identified the key priorities for modern banks:

(1) Develop a customer-centric business model.

(2) Optimizing distribution

(3) Simplifying business and operational models

(4) Gain information advantage

(5) Supporting innovation and the capabilities needed to foster it

(6) Proactively manage risk, regulation and capital

Innovation will be the main and most important factor for achieving growth in the banking sector in the next five years. This may take the form of a new product introduction, a change or introduction of a new process or a new customer experience, and at a global level for an organization it may even take the form of a complete transformation of its business model from its previous state. In PwC's survey of senior banking executives, innovation was rated as particularly important by 87% of respondents, and surprisingly only 11% of respondents admitted to being ready for it (Bob Sullivan, 2014). PwC suggest that digitalization will bring about huge changes leading to a reduction in revenue, the emergence of new competitors, redefinition of services and a transformation of the entire industry. Digitalization, new technologies and innovation are clearly the leading causes of global business change. Big data, cloud services, the increasing use of smartphones and broadband are factors that will transform the way personal services are used. The prediction, according to a case study by Uppsala University (Ortstad & Sonono, 2017), is that in a market like the US, digital transformation will allow regional banks to achieve national coverage without significantly increasing the cost of opening branches, and that developed international banks with a strong communications strategy will be able to benefit even more effectively from globalization. With smartphone penetration at over 65% globally (GSM Association, 2020), smart devices are already becoming the primary means of communication on the network. Their use in cashless payments and transactions is increasing, with a key challenge for providers being the identification of the correct user at any point in time to manage both the customer experience, the risk and likelihood of potential fraud and money laundering.

Technological innovations typically happen initially outside the traditional banking services sector due to various features: They require agile skills, minimal bureaucracy, lower dependency on management and shareholders, and freedom to think and experiment. For this reason, in order to create best practices and achieve a significant level of technological progress, banks can create an environment for innovation by partnering with different organizations: technology start-ups, academic or non-financial institutions.

2.3. Until a few years ago, savings, credit and payment products were offered only by banks, but now there are alternative providers for each of them, gaining popularity especially among individual customers: the physical networks of fast credit companies are even better developed than those of banks, some of them even financing the acquisition of a vehicle or real estate; online

payment platforms like Paysera attract customers with significantly more convenient and affordable services, and online trading and group investment platforms These, and similar examples of alternative providers, can be summed up in one word: fintech, a term that is increasingly being used to summarize a variety of alternative technological activities to banking in the financial world. Fintech companies have their role as a complement or alternative to the majority of products and services offered by traditional financial institutions. There are several areas in which they operate successfully. Beyond activities related to digitization and payment facilitation, in recent years fintech has successfully entered the core business of banking institutions in attracting and lending funds through so-called Peer-2-Peer (P2P) shared finance platforms. P2P is a concept that offers a direct linking of individual borrowers and lenders, achieving a more socially advantageous form of financing without the intermediation of banks and/or other conventional financial intermediaries. Banks are gradually starting to make digital progress, mainly by partnering with fintech companies, which are currently responsible for the biggest changes in the financial business in recent years. The credibility of partnerships with these types of companies is evidenced by venture capital funding support, amounting to nearly USD 44 billion since 2014, for fintech startups (Mastercard, 2018). In this context, traditional banking institutions need to define their development strategies in relation to alternative providers. Yet the balance of liquidity in the co-existence of loan, deposit and payments make the business model of a banking institution a unique enterprise (Milne & Parboteeah, 2016). At the same time, P2P platforms have the ability and potential to expand access to credit for smaller companies, high-risk startups, and serve individuals financially excluded from the formal banking system. Thus, accruing various benefits, both private, social and public, banks are obliged to find the right model to manage benefits and risk in a way that successfully combines and administers the various business and regulatory challenges in the digitalization of the financial sector.

3. Product diversification as an opportunity for customer capital growth

3.1. The success of any organization depends to the greatest extent on the business strategy it chooses and its ability to adapt and change it in the context of certain circumstances and changes in the environment. Long-term success is mainly due to a skillful combination of right investments, innovative thinking, creation of useful value and last but not least, a strong focus on strategic management (Colin Gilligan, 2009). The business philosophy oriented towards marketing

strategies as a means to achieve company goals has undergone a rapid development over the past and present century, undergoing several transformations. Recent changes in marketing concepts also reflect the significant role of technological developments contributing most to the personalization of marketing activities. The marketing planning of the company is necessary to have a vision of possible future changes in the environment and their effects on its development. Technological development and digitalization are dynamically changing the business conjuncture, and for many organizations this change is even unpredictable, risky and threatening. The objectives of strategic management include the definition of those competitive advantages on the basis of which the business can develop successfully. Competitive advantages are a consequence of the core competencies that the organization possesses. Based on the latter, the organization models its business strategy, the most popular model for possible strategic directions is the Ansoff Matrix (David Campbell, 2002) with four possible strategies. Applying it in the context of the telecommunications business, companies in the industry have successfully implemented the three options of market penetration, market and product development (excluding diversification). The challenge they face is reaching the limit of using these strategies as a means to achieve significant real growth and exploring diversification opportunities. Achieving growth through the lens of linked (organic) or unlinked growth is another cut representing possible directions of development (David Campbell, 2002). Improving performance based on an organization's inherent and acquired competencies for its industry is the foundation of linked business growth. A successful and difficult to imitate realization of disconnected growth would be possible when the organization finds the element of its core competencies that provide a meaningful basis for its competitive advantages. Diversification and unbundled growth would be successful when the firm's assets can be readily repurposed and the incremental costs are minimal.

3.2. Diversification is basically seen as product diversification - in the sense of a strategy of specialization in the existing business, related or unrelated to the core business of the organization. In recent decades, research on the relationship between diversification and performance has been strongly influenced by the Resource-Based View (RBV) (Diana Benito-Osorio, 2012). Its postulate is the understanding that the successful implementation of a diversification strategy depends mainly on the firm's available resources and capabilities. The RBV concept emphasizes the drive of firms to maximize their performance by making optimal use of available internal resources and capabilities in different aspects of the business. Companies across industries are investing in their

analytical capabilities and constructing databases of significant size and complexity to leverage the opportunities provided by digitization to create value-added competitive advantage. In their study (Noble, 2019), the authors examine the theoretical basis of the applicability of big data in the strategy selection process from the perspectives of the resource-based view (RBV) and organizational learning theory. By adhering to RBV principles and considering the heterogeneous nature of resources in relation to competition, companies' databases can be viewed as a valuable, rare, difficult-to-imitate and substitute resource, which automatically makes them a potential competitive advantage. Their power and utility is amplified when a company has the capacity and capability to extract from the data and put into action in different business models, knowledge and insights established in a particular way. Companies following the RBV approach are identified as those that focus on improving their existing capabilities to build internal infrastructure and create skills to solve existing problems. In contrast, firms with an organizational learning orientation are focused on seeking new products or non-traditional market opportunities. In this sense, it makes sense for companies following the RBV approach to outsource some of the analytics capabilities to an external entity or form a strategic partnership with an organization with sufficient competencies in solving big data business problems. The telecommunications sector, as the largest area of public and business life receiving, storing and processing big data, has the potential to use it not only for its own needs but also to generate significant financial flows from its use for alternative purposes to its core business.

3.3. From a business point of view, it is a valid statement that the only sustainable competitive advantage comes from out-innovating the competition (Colin Gilligan, 2009). In contrast to traditional business models that aim directly at profit maximization, the modern concept of CRM (Customer Relationship Management) relies on creating, developing and maintaining long-term relationships with customers, believing that these are more profitable and effective (Morgan, 1994). The CRM approach modifies the customer value theory and accordingly is invariably related to both customer lifetime value (CLV) and customer equity (CE). When calculating CLV, the profit margin of an individual customer has the largest effect on the result due to its multiplier effect by the number of users, especially in the presence of a large customer base. Therefore, the telecom and banking industries always seek to upgrade their customers with as many additional services as possible, increasing the average revenue per subscriber. CE-led companies see themselves not as a portfolio of products but as a sum of customers and each customer segment or sub-segment is

considered a separate source of revenue requiring to be managed efficiently and individually. Accordingly, maximizing customer equity is achieved by maximizing the lifetime value of each individual customer, group or market segment (Francis Arthur Buttle, 2015). To successfully manage customers and view them as a company asset, customer data needs to be structured in a dedicated data warehouse (DWH). Storing customer data sets, organized in a way that allows them to be processed with appropriate tools for complex analysis, helps to reduce marketing costs by simultaneously applying individual offers and selected information to large groups of customers united by common characteristics and behaviors. By using intelligent data analysis instead of observation, companies can more successfully predict future actions and purchase probabilities even at the individual customer level and thus enable the transformation from mass through niche to individual marketing.

3.4. Information and communication technologies (ICT) are continuously creating new types of market segments and new patterns in business development. This refers to the way in which actors in a given business area interact or compete in the value chain. Through convergence, ICT companies make it possible for organizations from different industries that have never been connected to compete or collaborate. Diversification is a common idea through which such cross-industry interactions are made possible. For telecom and banking institutions, development and upgrading through information technology is a must, and mobile banking is for now the link between the two, following the advent of internet banking. In terms of other financial services, the major global ICT companies (Google, Apple, Facebook, Samsung, etc.) in the mobile supply chain are already attracted by the potential and are providing digital-first offerings for customer payments. However, the advantage telecom operators have over them is that they (telecoms) ultimately have the ultimate control over consumers and are in constant contact with them. Telecoms and banks realize the opportunity to both collaborate and become competitors. Therefore, there is mutual distrust between them (Sangjo Oh, 2006). A fruitful partnership between the two industries can be based on strong mutual support and excellent communication. Good coordination in partnership relationships could be based on strictly defined roles, responsibilities, rights and obligations. Possible joint business models imply different roles and functions of the participants. Strategic cross-industry partnerships have not yet materialized, but the growing need for integration of financial and mobile services will drive such developments. Time will tell whether partnerships will evolve into mergers or into joint service portfolio offerings, with the possibility of telecom to

banking transformation or vice versa through one sector entering the other. What can be said with confidence is that the two industries, which had nothing in common 10 years ago, will become increasingly linked and interdependent.

Chapter Two: The provision of financial services and credit risk

Chapter Two deals mainly with the issues in credit risk assessment. It provides an overview of the main data and indicators forming credit assessment methods, their advantages and disadvantages. The alternative data that the society generates in the present century, their original use in different spheres of economic activity are considered and the proposal is formed they are used to improve the models of credit risk assessment.

1. Credit risk assessment

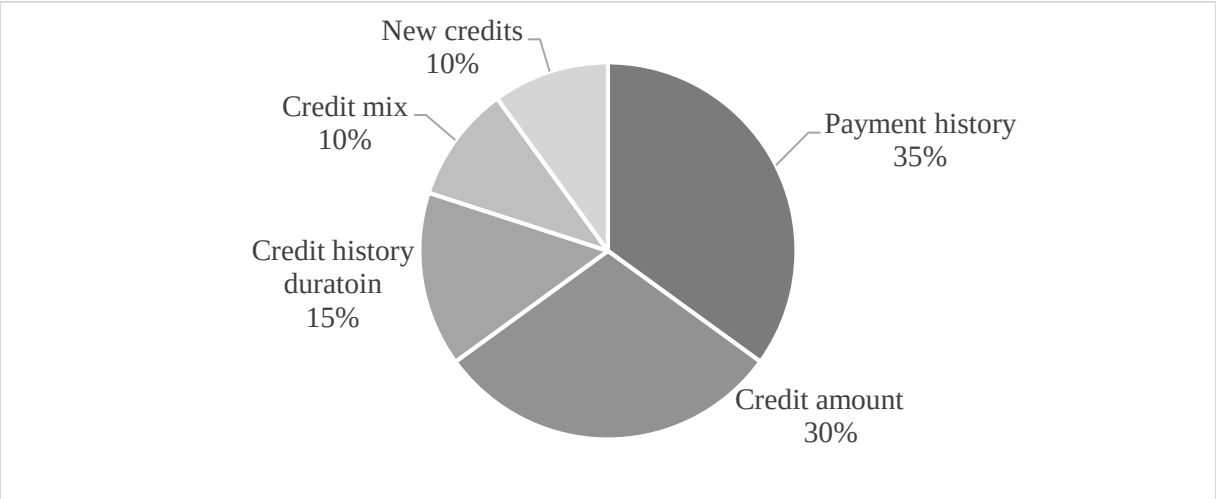
1.1. The main services in the business model of financial institutions are financial intermediation, the raising of finance through deposit-taking or wholesale funding (through bond issues) and/or equity and lending, the latter being a credit risk assessment activity. A major challenge for lenders is the need to assess the creditworthiness of an applicant. Financial capacity is generally based on income information from an employer. Situations where a customer has credit positions with more than one financial company and desires an additional loan led to the emergence of credit or rating agencies or bureaus in the 1980s with the primary purpose of collecting and processing customer information regarding the consumer's obligations to various institutions. This data was arranged to be made available, when necessary, to service providers related to potential financial risk.

The key players in credit reporting systems are several: The data sources are those that must provide correct, timely and comprehensive data on a fair basis. Other data sources must facilitate credit report providers' access to their databases; Credit report providers ensure secure and quality data processing and efficiency of the services they provide; Users who use the information provided to them correctly; Data subjects must provide reliable and accurate information to other participants in the process.

According to the World Bank rules (IBRD/The World bank, 2011), regulators should encourage the development and use of credit reporting systems because of their effectiveness and efficiency in meeting the needs of different actors in the financial system; respect the user rights of data

subjects and support the development of a fair and competitive credit market. The credit information collected and processed is of substantial interest for the successful business development of various market participants. Specialized units known as credit reporting service providers (CRSPs) act as intermediaries for credit data users, collecting it from individual and business customers. They can be divided into two types - credit registers and credit bureau. The main role of credit bureau is to improve the quality and availability of data in order to enable financial and non-financial creditors and other users to make better informed decisions. For credit registries, the priority is to participate in banking supervision and improve the quality and availability of data on supervised financial intermediaries. Globally, coverage of both types of institutions is in the order of 20%, which is unsatisfactory (The World Bank Group, 2020). As a result, there are currently large groups of consumers in some regions who are excluded from their formal financial systems, a huge potential for providers who, given an alternative assessment system, could substantially develop the supply of financial and credit services.

1.2. A customer's individual credit score is the indicator of the likelihood of default or inability to service a loan. It is a numerical score based on a calculation of various indicators, mainly related to the individual's credit history, weighted by a certain weight. In the most general case, a period between 12- and 24-months back is analyzed for new customers, and 6 to 18 months back for existing customers. In recent years, more and more financial institutions have built complex and sophisticated models that predict not only the probability of non-payment but also the propensity to fraud, bankruptcy or over-indebtedness (Scott, 2002). The factors involved in calculating a credit score are developed from the historical data available in a customer's credit file at a repository, agency, or other source. The number of points received in the rating often determines whether the applicant can obtain credit, the maximum amount and its cost (Experian, 2020). A world leader in creating credit scoring models is FICO, an abbreviation for Fair Isaac & Co. The general structure and approximate weights of the metrics included in the credit score calculated by FICO and its competitors is presented in the following Figure 1:



Source: (Akin, 2020)

Figure 1: Structure of indicators involved in credit scoring

A credit rating or score has the same basic purpose: to assist the lender with information about the applicant's associated risk of loan default. A high rating is logically associated with a low probability of default, and a low rating indicating poor debt management serves as a reason for lenders to avoid such applicants. But new credit applicants also have relatively low scores due to a lack of credit history. Therefore, lenders often use credit ratings along with other information such as employment history, available savings, other assets, or evidence of income to decide whether they are willing to provide credit and at what cost (Akin, 2020).

1.3. Alternative data, along with historical financial data, a new category of personal information is emerging, used for a variety of purposes: modern technologies, including machine self-learning and artificial intelligence, are able to determine in real time what information we should be shown online, what our needs are based on where we live or work, and even how to vote based on our personal disposition, among many others. The most famous example in the last decade of using behavioral data to determine 'consumer' preference was Donald Trump's campaign for 2016 US presidential election. Company Cambridge Analytica was accused of having improperly used personal data to influence the vote in the election. The "Alamo Project" matched OCEAN user psychological profiles (openness, conscientiousness, extraversion, agreeableness, and neuroticism) with social media information and "news" available on the Internet, with the primary task of identifying and canvassing voters who would vote for Donald Trump and dissuading those who planned to choose another candidate. The basis for the data is an analysis of 2010 election results.

Another large-scale project based on data generated in the space is the Social Credit System (SCS) in the PRC. The model links public and private data of financial and social behavior of individuals and organizations in China, and the results are used to obtain a score: reward, reward or punishment defined according to certain agreed standards of appropriate behavior (Martin Chorzempa, 2018). The SCS has two main components: (i) to become the world's largest data set, integrating currently disconnected data from government and non-government organizations in China and expanding its efforts to collect and warehouse behavioral data from a variety of sources. (ii) The second component in a system is the carrot and stick concept to make people and organizations more "sincere" and "trustworthy".

Less spectacular but successful projects using behavioral data for prediction can also be found in the field of telecommunications. Gender Analysis and Identification Tool -Gender Analysis and Identification Tool - GAIT is a software for predicting the gender of telecommunication service customers based on their CDR records using machine self-learning (Dalberg Group, 2018). Such examples illustrate the variety of applications of behavioral data for alternative marketing activities and demonstrate the power and potential of machine learning algorithms to discover and exploit the hidden information behind customer characteristics and actions and their applications in theory and practice.

2. Credit risk assessment based on telecom data using CRISP-DM methodology

2.1. Big data analysis and processing skills are known by the term Data Mining (DM). It is a process of creativity requiring a variety of knowledge and skills. As noted in the publication (Hipp, 2000), it consists of the ability to transform a specific business process into an intelligent analysis problem to be solved by specific data processing using statistical and computational techniques that ensure the efficiency of the result. The adoption of a common working model in the Data mining process is of great benefit to analysts. The generally accepted model used by professionals for knowledge extraction from large databases is known by the English acronym CRISP-DM (Cross-industry standard process for data mining) or "industry standard process for knowledge extraction from data". Data Mining projects are complex tasks with high expectations and therefore need to follow a clear methodological framework and specific steps outlined in it. The CRISP-DM methodology divides the Data Mining project lifecycle into six phases and associated specific tasks, illustrated in the Table 1:

Table 1: Detailed CRISP-DM process and related tasks

Business understanding	Data understanding	Data preparation	Modeling	Evaluation	Deployment
• Determine business obejtives • Assess situation • Determine DM goals • Produce project plan	• Collect initial data • Describe data • Explore data • Verify data quality	• Select data • Clean data • Construct data • Integrate data • Format data	• Select modeling techniques • Generate test design • Build model • Assess model	• Evaluate results • Review process • Determine next steps	• Plan deployment • Plan monitoring and maintenance • Produce final report • Review report

Source: (Hipp, 2000)

The combination of in-depth business knowledge along with the power of Data mining modeling can realize successful projects that contribute significantly to forming a competitive advantage of the company in terms of management and customer satisfaction.

2.2. In this thesis, the possibility of predicting individual income of telecommunication service customers based on socio-demographic and behavioral data will be tested using CRISP-DM methodology. For this purpose, an anonymized database of 110,000 individuals subscribing to a mobile phone service is used. Due to the nature of the study, the dataset examined here contains only data on individual customers, and only includes those with one SIM card registered in their name. The following types of data are available: Demographic- age, gender and location; Static telecom data- year of registration (indicates first registration of SIM card), tariffs and tariff group, tariff type (internet included or not); Dynamic telecom data- traffic data are averaged and are minutes, SMS, MB internet usage, international calls; financial data- declared income (in case of change of place of work, non-declared data, long sick leave, unpaid leave, the data are partial); calculated data- share of minutes of calls to family and friends, total number of minutes of calls, share of minutes to other national networks in total outgoing minutes. For the purpose of applying classification analysis, a variable indicating the declared income below (by 0) or above (by 1) the

national minimum wage has been added. Since the minimum wage is a key indicator of the subsistence minimum, for the purposes of the analysis it is assumed to define the threshold below which a client is considered to be at risk and most likely to be unable to service a potential financial obligation.

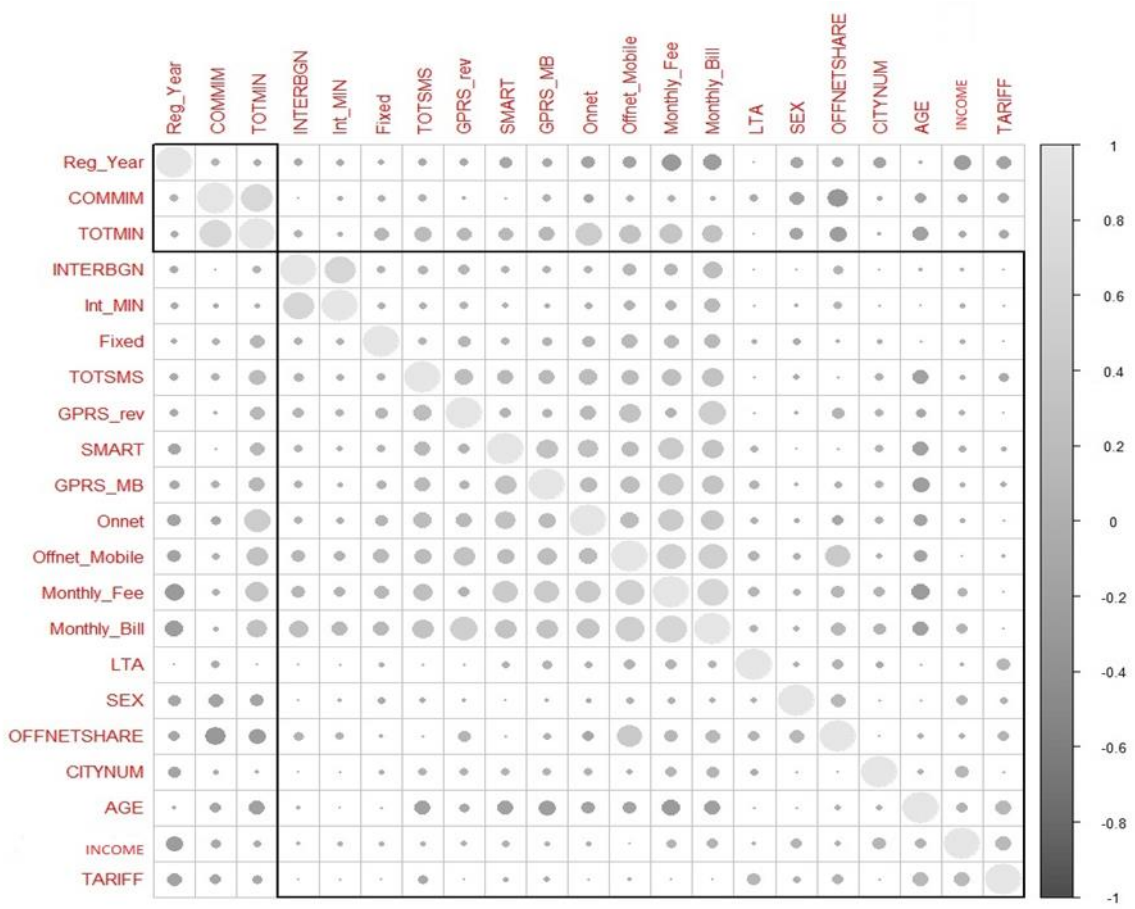
2.3. A preliminary view of the database tested in the development is provided in the subsequent graphs. Table 2 provides an overview of the main descriptive statistics for the monthly data, with a full set of data available for all sample participants i.e. there are no missing values for any customer:

Table 2: Descriptive statistics of monthly data

		Age	Income	Registration year	Monthly fee	Monthly bill
Mean		50,51	6341,24	2005,99	11,32	16,65
Median		50,00	4552,33	2007,00	9,11	12,60
Mode		43,00	2880,00	2008,00	8,25	8,25
Std. Deviation		12,36	5531,54	3,00	7,02	13,61
Skewness		,11	1,61	-,84	1,77	7,58
Kurtosis		-,88	2,81	-,26	5,63	182,53
Percentiles	25	41,00	2781,00	2004,00	5,75	9,03
	50	50,00	4552,33	2007,00	9,11	12,60
	75	60,00	8581,92	2008,00	14,71	21,10

Source: Author's calculations

In addition Figure 2 below is a correlation matrix of the variables presented. It clearly distinguishes the existence of a relationship between the total number of calls, the monthly bill and charge and the rest of the calls. The type of tariff plan (Smart), mobile data usage (GPRS MB) but not its revenue (GPRS_rev), and the monthly charge and bill are positively correlated with each other and negatively correlated with the age indicator, showing that mostly young people move towards tariffs with mobile data included (which are not paid separately, so there is no change in GPRS_rev) with higher charges and therefore monthly bills.



Source: Author's calculations

Figure 2: Correlation matrix

The strongest correlation is observed between the year of registration, the field area and the group of tariff plan used. Applying predictive models through analytical knowledge extraction may reveal additional dependencies not reflected in the matrix shown.

3. Compilation and application of predictive models for credit risk assessment in a telecom

3.1. The main problems and challenges in defining/calculating credit score are the insufficient financial history of the potential customer and the possibility of committing financial fraud. In the absence of historical data, the risk to the financial institution is considered high and customers receive a low score. However, behavioral data on consumers themselves exist in many other non-financial records and their use in the context of financial services consumption would facilitate and refine individual credit assessment. This thesis aims to test and demonstrate the feasibility of using

such data for financial purposes. Intuitively, a positive correlation between the use of mobile phone services and a consumer's financial income could be assumed:

- Monthly subscriber usage data generated by telecom companies can be used as alternative data to predict customers' financial income
- Based on the results obtained from behavioral data predictions, a telecom company can estimate the financial reliability of each of its customers
- Analyses of big telecom data, in addition to core business, have possible applications for purposes in other areas and fields of economic activity
- Telecom companies can move beyond their role as a mere platform for providing mobile financial services and become an active player in the market for borrowing funds from individual consumers

The following chapter will test the possibility of classifying customers conditionally as "reliable" (those with an annual declared income above the minimum wage) and "unreliable"-customers with an income below the minimum wage. The classic confusion matrix will be used to visualize the result. The approach that aggregates the sensitivity and specificity indicators for all possible threshold values is the so-called ROC (Receiver Operating Characteristic). In traditional scoring based on historical data, the classification task separates a threshold separating "good" from "bad" customers based on past misconduct, trying to find similarities between the performance of a credit applicant and similar performance of a customer with whom the financial institution already has some experience. In this sense, even if alternative telecom data predicts personal income with relative accuracy, it (income) does not guarantee the correctness and predictability of financial behavior.

Financial institutions have data on the education, employment, and ownership of their customers, which can also play an important role in a cross-industry study. Such data are not available in the database provided here, although their presence would have improved the outcome of the study. It is important to note that the data used in this study are for individuals declared insurable income i.e. non-income, including from regulated or non-regulated activities is not available. According to AICB data (www.mediapool.bg, 2020), the share of the informal economy in Bulgaria amounts to 21%, which represents a significant difference between declared and real income and has a corresponding impact on the results of this study.

3.2. The process of implementation of the set research task begins with the correct construction of the Data Mining study, determining the appropriate methods for the problem, corresponding to the structure of the sample and the indicators available in it. By its nature, the objective of telecom and financial data analysis in this study is to predict the values of the outcome/dependent variable (declared income) based on past values of the independent variables and their dynamics (demographic characteristics and telecom consumption). For this purpose, the models to be tested are classification and regression analysis. The model used is of the Self-learning with teacher type where an algorithm is constructed using part of the available data to predict the value of a dependent variable using inputs called predictors. The "trained" algorithm is then tested/validated on the rest of the basis and a comparison is made between the resulting calculations and the actual values of the dependent variable (Иванов, 2016). Algorithms used suitable for a numerical, continuous dependent variable such as individual income are: classification and regression decision trees and neural networks. When solving a classification problem where the outcome variable is categorical with two alternative values- "approved" and "rejected", classification algorithms are used where the predicted dependent variable is a dichotomous variable: some of the most common methods tested here are CHAID, CART, C5.0.

Also used in the development are the matched (ensemble) methods (Rocca, 2019) Bagging and Boosting, in which multiple "weak learner" models are trained to solve the same problem and then combined/integrated to obtain a better result. The main goal is to combine weak models correctly to produce an improved model.

3.3. The benefits of applying alternative data for financial valuation purposes are of great importance for successful financial service delivery. Despite the reliability of the currently used data and regulatory restrictions, according to a study by Global Metrics (Глобал Метрикс, 2018), even in the EU the average percentage of households with arrears in 2016 was 10%, and in the Balkans: in Greece it was 47%, in the Republic of N. Macedonia- 42%, Serbia and Bulgaria with 37% and 34% respectively. The above-mentioned shortcomings in modern lending can be somewhat eliminated by enriching the credit assessment with alternative telecom data. Customers of fast credit firms could be segmented in more detail, with a fair risk premium and a linked probability of default for each segment. Telecom data is not only related to the customer's financial past, it is as up-to-date as possible, reflecting current customer behavior and could be used in the

context of predicting future service user behavior. In addition to assessing the risk of default and non-payment, another problematic aspect of the credit scoring methods currently used, as shown by research in the US (Johnson, 2020) is the worrying trend of high rejection rates when applying for credit. A 2013 study (Yu, McLaughlin, & Levy, 2014) found that 20% of traditional credit reports contain errors, and in 5% of reports, the errors result in a lower credit score that can make credit more expensive or impossible. Having alternative data to validate and supplement credit scores would lead to more thorough checks that identify and correct inaccuracies early on.

Chapter Three: The Role of Big Data Analytics in Diversification as a Growth Strategy for Telecommunications Companies

This part of the study presents the result of analyzing behavioral telecom data and testing machine self-learning models to predict individual income. The benefits of applying these models in practice are presented, as well as the market potential for mobile financial services development. Possible strategies for telecom and banking convergence in financial services are reviewed and a specific business model proposal is demonstrated.

1. Predictive analysis using behavioral data

1.1. Data mining techniques and models were applied using several machine learning software products IBM SPSS Modeler[3], JMP[4] and Orange[5] on the provided telecom customer database to predict the declared income of the individuals in the sample based on the demographic data available on them and their mobile phone service consumption.

The different research models show the best results when tested on the training sample, and when considering the RMSE indicator, the algorithms applied by Orange give significantly better results than the other models, and the high value for the linear correlation method should indicate the non-linear relationship between the variables. When ranking the indicators with the greatest influence on the outcome variable, the predictors with the greatest importance in the CART model are the year of registration and the age of the customer (logical due to high income subscribers having purchased the service earlier), followed by other demographic characteristics such as birthplace

[3] https://www.ibm.com/products/spss-modeler (10.09.2021)
[4] www.jmp.com (10.09.2021)
[5] https://orange3.readthedocs.io/projects/orange-data-mining-library/en/latest/reference/evaluation.cd.html (10.09.2021)

(city) and gender. Of the telecom data, the most important are the group of tariff plans and the number of SMS sent per month. In the first division in CART, the model splits customers into those with registration before and after 2005, and for those existing before 2005, the second most important metric is tariff plan, followed by age, with customers older than 34 subsequently split by SMS consumption. For subscribers registered after 2005, age is the first most important criterion, followed by location, tariff type and gender. The ranking of the importance of the predictors of Neural Network usage is the data starts from the telecom consumption data - the monthly bill and the amount of monthly calls in minutes are first and second, followed by the age of the customer. The next four places are again telecom data, followed by the year of SIM card registration.

The patterns in the different software products show similar results, which finds a logical explanation from practice: the main predictors are time-of-use related characteristics - year of registration, age and tariff group (different tariff groups were available at different time periods), followed by monthly bills and total consumption. Despite the wide range of results, given the limitations of the dataset, the models were able to show with relative accuracy the existing relationship between declared individual income and available demographic and telecom usage data.

1.2. Predicting a continuous numerical quantity such as declared income from telecom consumption data is a task whose accuracy is relative, and the measurement and estimation can be largely described as subjective. Although income prediction algorithms can be applied, they are not able to estimate the probability of fulfilling and partially or fully defaulting on a financial obligation. Similar to the way financial institutions classify their customers with alternative coefficients: reliable (denoted by 1 - ts as those who can be granted a loan) or unreliable (denoted by 0, respectively), the study becomes a classification problem with a dichotomous dependent variable: A threshold is established, in the amount of the Minimum Wage, which conditionally divides the clients into two groups: in case the annual declared income of the client is greater than the annual minimum wage (AWW), the dependent variable will have the value 1, and in the opposite case - 0 (zero). Of course, the threshold at which the decision is made can be programmed at different levels, but in this case, for the purpose of the study, we need to justify a product that is not competitive with bank credit, but is its complement. Due to the nature of their policies and the strict rules and regulations to which banks are subject, their requirements of potential customers

are very strict, which has resulted in an ever-increasing niche market occupied by so-called quick credit firms, satisfying financing needs with loans that are minimal in amount, serving risky customers. The logic in this problem is to assume for the purpose of the study that, with an annual income below the minimum, the potential customer will be in difficulty or unable to meet his financial obligations and, therefore, the associated risk will be excessive. However, in order to reach a larger number of clients for whom the repayment of a micro-credit, which will be used for exceptional one-off needs and is strictly limited in time, with a maximum repayment period of up to a year, the existence of any reasonable declared income is sufficient justification for its granting.

For all the classification algorithms used in the study, the measured AUC was above 0.65, with the highest performing methods from the JMP software product, with the Bootstrap Forest method offering the most accurate prediction of AUC with levels near and above 0.7. It can certainly be argued that the data and processing used can be successfully applied for forecasting purposes, however it is insufficient in its current form to service a larger number of potential customers. To be used on their own for credit risk assessment at an AUC level below 0.7, the number of at-risk customers who would receive financing but be unable to repay would be too large. In order to compensate for this discrepancy, the cost of funding would need to be at a level that would cover the potential losses of the 'bad' customers, which excludes the mass condition. The successful existence of a sustainable and popular funding model is directly dependent on the 'risk premium' condition (the interest rate) and must be tolerable. This premium is a function of the accuracy of the measurement of individual risk: the lower it is, the smaller the risk component (interest) will be and at the same time the number of customers will be maximized. In order to achieve higher accuracy in forecasting algorithms in the order of AUC 0.85-0.9, data with a higher degree of confidence and calculations with a more detailed telecom data set (e.g. not average monthly amounts, but detailed tables at the level of each individual event) are needed. On the other hand, the share of the informal economy in the country is too high and the official income mentioned in a document could not be the only guarantee of creditworthiness. In addition, it is possible to examine, compare and analyze the different types of customers in the telecom customer database: In this case, only customers with a service contract with their telecom provider are included here. Prepaid service users are not present in the provided database. In this sense, in order to correctly predict individual income, it is necessary to compare the type of subscribers with a contract and a prepaid service and to confirm or reject the logical lifetime thesis that the former have a contract

due to the security and predictability of the income they receive (and as such are almost all low risk) and the customers with low or no income are those using a prepaid service.

1.3. Financial institutions, over the long life of their existence, have built up strict credit scoring rules that are constantly being refined and built upon. Strict banking regulations require maximum protection for creditors, but this happens in some cases even at the expense of bona fide customers. The rise of MFIs proves that there is a niche market with enormous potential, which should be explored in detail and exploited through alternative means, due to the lack of sufficient flexibility on the part of banks to serve it and the lack of MFIs of the power of banking institutions, their processes and information technology to fill the absence of models and alternatives necessary for the financial inclusion of risky client groups.

The success of the Upstart business model[6] in the US can be cited to support the above assertion. It considers the future employability and potential future income of borrowers, based on their education, the institution they attended, their major and their academic performance, instead of analyzing past (historical) data. Based on these inputs, the company develops a model that calculates the probability of default on the loan for each potential customer. Another example of successfully applying statistical models on telecommunications data to achieve non-telecommunications outcomes is the Cignifi[7] credit score. The company's desire is to demonstrate the feasibility of exploiting mobile usage databases to identify potential customers with savings propensities. The pilot was implemented in Ghana in 2013 with data on 1,600 customers with accounts at the local Airtel telecom and the local HFC bank. As a result, five customer segments were identified, based on CDR analysis and behavior in different previous marketing campaigns and different communication channels used by both the bank and the telecom (Hakim, 2020). The number of such examples is growing over time, which should show the huge unexploited potential of such initiatives based on the exploration of hidden information in utility providers' datasets.

2. Telecom financial institution

2.1. The potential of microfinance services in Bulgaria and the rest of the world is enormous, as evidenced by the statistics available in the BNB files[8]: at the end of December 2014, the share of

[6] www.upstart.com (10.09.2021 г.)
[7] https://www.cignifi.com/ (28.10.2021 г.)
[8] https://www.bnb.bg/Statistics (10.09.2021 г.)

loans up to BGN 1,000 was 47.8% of the total (vs. 49.5% at the end of 2020), and at the end of 2009 it was 47.3%. However, the share of small loans is decreasing as a share of the total loan portfolio of banks. The average size of loans to households at the end of 2020 amounted to BGN 9 194. In comparison, at the end of 2019 it was BGN 7,909 and 5 and 10 years earlier (2014 and 2009) it was BGN 6,840 and BGN 6,651, respectively. The number of non-bank financial institutions registered with the BNB at the end of 2020 was 200, compared to 184 at the end of 2017, The amounts of loans granted by them have been growing on average by 8% on an annual basis for the last 5 years and exceed BGN 3 billion at the end of 2020, which is almost 50% more than the amounts of loans granted by banks in the segments up to BGN 5 000. The preferences associated with MFIs are confirmed in a study on non-bank financial services conducted by Nielsen Admosphere Bulgaria (Nielsen Admosphere Bulgaria, 2020) in December 2019. According to it, about 63% of respondents between 35 and 44 years have taken out a loan from a non-bank financial institution at some point in their lives, while in the remaining age groups the indicator varies between 44% and 49%. Advantages of this type of financial service cited by respondents were the ability to withdraw a small amount of money in a short period of time (cited by 38% of respondents), the ease of obtaining credit and the quick negotiation and ability to apply at any time (important conditions for 32% of respondents). At the global level, the potential for the development of alternative data-driven financial services is significant and still undeveloped. As stated in the Global Financial Database (Demirgüç-Kunt, Klapper, Singer, Ansar, & Hess, 2017), the estimate as of 2017 is that 1.7 billion people in the world are disconnected from formal financial systems - they do not have financial accounts with either traditional financial institutions or mobile wallet providers. In developed countries, almost all individuals have a bank account, formal or virtual, which implies that the proportion of those financially excluded is significantly high in the developing world. Individuals who do not have access to formal banking, i.e. who do not have a financial account and cannot be served by one, including those who do not have access even to microcredit, which is particularly important for ensuring a minimum standard of living or starting a business, can be served by MFIs. Micro-lending based on alternative risk assessment for individual clients without financial history is the next logical step in the process of developing MFIs to offer a full range of services to the financially excluded. It is no coincidence that these businesses are behind the World Bank's initiative's main goal of 1 billion people without financial accounts acquiring one by the end of 2020 (Worldbank Group, 2018).

2.2 Mobile services: globally, the financial services sector encompasses a large number of products and services, and its development by region and country is at varying degrees of maturity. Mobile financial services (MFS) have a number of advantages for customers: in addition to being quick and easy to use, they save time and create convenience. Their use saves the need to be physically present in branches and other institutions. The development and distribution of MFS is not an exclusive priority of companies in the financial sector. For a variety of subjective and objective reasons, a significant number of non-financial companies, such as retailers, technology platforms, utility companies or automotive ones are developing their own high-tech financial services offerings. Often such non-financial companies enter the financial sphere having already experienced their own transformation through the introduction of innovative technologies. They have the experience and knowledge to redesign initial customer propositions in a way that makes them faster, cheaper and more convenient. A start-up company set up by a retail chain or utility company, for example, would have no problems with external funding and could benefit from the flexibility of the newly created structure under its control. At the same time, customer preferences for financial service providers are not limited to traditional ones, despite the insufficient level of trust in non-financial companies. Their preferences are highest for retailers (45%) and telecommunications operators (44%), followed by technology companies and social media (41%) (Bull, Chen, & Chiselita, 2019).

2.3. The existing cyber risk to consumer data acquired and used by financial service providers, whether financial institutions or not, should not be underestimated. Platforms such as YouTube and Instagram, with over 1 billion active users, sell their customer information to brands that use it for advertising purposes on social networks. In such an environment, telecoms and fintech companies, despite their differences: industry, scale and culture, have ample reason to join forces to provide innovative, secure and useful services that perfectly complement each other based on scale and know-how. The benefits of cross-sector, including telecom-finance partnerships have already been identified as possible digital growth strategies by the leadership of major telecom groups globally. In a collaboration between a financial institution (FI) and a telecom, the FI would benefit from the following advantages (GSM Association, 2017):

- The telecom's mobile portfolio and its integration with the financial institutions’ services
- Access to a major source of big data and its analytics

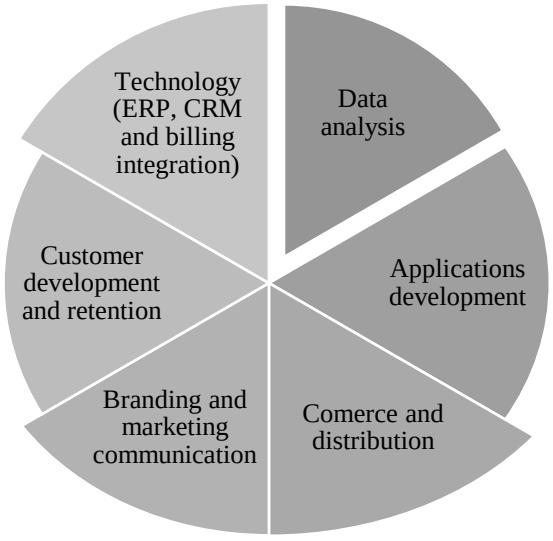
- Potential opportunity to increase its customer base through more effective retention and lower churn rates.

Telecoms would have the following benefits:

- Broad utilization of their telecommunications network
- Access to multiple financial instruments and products from the banking sector
- Additional revenues from bundled services (including financial services) available in mobile wallets

3. Potential business model

3.1. Clearly, both business models operate in a regulated environment and, unlike the FMCG sector, have established long-term relationships with their customers and frequent interactions with them. With the trend of digitalization and the creation of a parallel online world, it is necessary for both industries to initiate actions for closer cooperation with each other. Cross-industry partnership optimizes companies' operations by reducing direct costs, reducing indirect costs through increased customer loyalty and consequently lower acquisition and retention costs, and increasing direct revenue per customer. The partnership between a telecom and a financial institution to offer MFS is characterized by the presence of a number of advantages. The two types of companies have significant, powerful and expensive infrastructure that can be shared, shown on Figure 3:



Source: (Baschnonga, и др., 2016)

Figure 3: Infrastructure that can also be shared by financial institutions and telecom companies

The companies have a set of assets and internal processes that can operate in close collaboration, which can have synergistic effects in both the existing business and the joint MFS offering. Shared use areas can not only contribute to the optimization of structures, but also in most cases complement the expertise of the partners on essential details of the response business.

The choice of a particular partner model should be made after an analysis of the factors influencing its eventual success: the levels of development of the telecom and financial regional markets, the development of technology, customer preferences and needs and, last but not least, the strategic objectives and the available or possible competences of the organization.

3.2. The main objective of bank risk management policies is to build technological and organizational systems to effectively manage the risk borne by the bank. One of the main means of risk management is risk prevention, and it is usually related to the bank's refusal to take credit risk, i.e. to refuse to lend and the potential income from it. Undoubtedly, telecoms are at the center of the mobile digital market and a coveted partner for banks and fintech companies as well as global technology leaders such as Google, Apple, Samsung, etc. in the mobile payments space. They all offer digital wallets and services with the help of telecom operators. With the number and volume of transactions constantly increasing, it is important for the latter to proactively increase their share of activity in this process. In their analysis, ACI Worldwide researchers (McDonald A. , 2017) point out that telecoms currently only manage to charge 3% of mobile finance-related revenues, with the market volume in Europe alone forecast to double between 2017 and 2023 from €114 billion to €231 billion. They therefore need to find innovative ways to increase their role in this process. A technology and information partnership between banks and telecom, based on data sharing and analysis, could successfully become a P2P platform for the provision of savings and credit products. Upgrading with services to

- Tailor solutions according to customer profile
- Cost control
- Opportunities for investors
- Provide knowledge to customers
- "Tailored offers"

Joint venture can outperform MFIs in offering short-term loans. Improved credit scores are likely to reduce the cost of borrowing capital, attracting more clients who currently use alternative sources of finance and refinancing. As Europeans become more mobile, the provision of cross-border services is becoming increasingly important. A joint initiative between a bank and a telecom could provide them efficiently, quickly and at negligible additional cost due to the experience and knowledge of international telecom operator groups.

3.3. To a large extent, the direction of future development of financial services can be traced and predicted on the basis of the processes taking place in Tropical Africa. In an effort to focus on the importance and potential of mobile financial services, some of the telecom companies in Tropical Africa are introducing the CDO (Chief Digital Officer) managerial position and even now expanding it to Chief Digital and Fintech Officer[9], which is also a reflection of the huge role and expectations towards the impact of mobile finance on the telecom business in the region. The potential of collaboration between a telecom operator and a financial enterprise can develop a mobile financial services platform serving micro-credit, savings and payment services. A business model for such a venture implemented through a stand-alone structure that maximizes the use of the participants' own assets would have the following distribution of roles:

Table 3: Business roles of participating parties in the proposed model

	Start-Up	*Telecom*	*Bank*
Product	●		
Marketing	●	●	
Sales		●	
Customer Service		●	
IT Integration	●	●	
Investor		●	●
Credit assessment		●	
Credit Risk		●	
Software	●	●	
Lender		●	●
Collection of non-performing loans			●

Source: author's concept

Expansion of the business of providing telecom related commodity credit, by outsourcing it to a 'start-up' structure, could provide the impetus for this business to be subsequently expanded to provide credit for various needs under a digital non-banking financial institution model.

[9] https://www.mtn.com/yolanda-cuba-to-join-mtn-as-group-chief-digital-and-fintech-officer/ (5.7.2021)

Considering the size of the market for such businesses, the innovative and fast business model proposed and the huge capacity of the companies behind it, the telecom-finance subsidiary has the potential to become one of the leaders in providing microfinance services to individual customers.

Conclusion

In summary, the following is an overview of the objectives of the dissertation:

- *The opportunity for telecommunication companies to add value by using the diversification strategy is demonstrated*. Offering financial sector services from a telecom subordinate structure is a viable and powerful growth strategy through an alternative service, leveraging the technology assets available in telecom primarily, used to satisfy the growing market potential of non-bank financial services.
- *Cross-industry leveraging of big data has a key role in the convergence between different areas of economic activity*. Behavioral data already has its uses in various business models, and its use in determining individual credit scores in the absence of sufficient historical financial history for the consumer has a direct impact on the cost of credit and financial inclusion in general for at-risk communities.
- *Predictive models are demonstrated to analyze behavioral telecom data and the existence of a relationship with consumer financial health*. Despite the limitations of the data used, a dependence of reported income on telecom consumption data is shown. The ability to use and analyze data at a level of granularity down to each individual transaction could increase the accuracy of the study, reinforcing the thesis on the benefits of using behavioral data to predict non-directly related metrics.
- *The possibility of customer segmentation and classification against minimum individual income based on behavioral and demographic data is very real and feasible*. Defining a low financial barrier for potential clients increases access to equitable financial resources for disadvantaged segments of society, which in addition to generating additional income is an indirect contributor to the development and growth of society. Micro-credit financial service with balanced risk, but with greater accessibility compared to that of traditional financial institutions, is a stimulus for economic activity of both the middle class and the poor and is an important condition for reducing social inequality in society.

- *A possible business model for offering online financial services is proposed, operating mainly with the active participation of the telecommunications company*. The potential of such a model is enormous, and the challenge for its construction is the mandatory definition of operational autonomy and independence, which would guarantee that the success of such a project would be measured by activity-specific indicators and would not be administratively and functionally controlled by the owning company, but at the same time would make maximum use of its available technical, financial, legal, human and expert resources.

III. Future research guidelines

Various studies reviewed in this dissertation propose the use of alternative behavioral data for a variety of business purposes, including the telecommunications industry. Investigating and analyzing the relationships between individual consumption of mobile services and unrelated metrics such as e.g. gender, age, income, employment, health status, etc. presents a relative difficulty due to the lack of sufficiently reliable data to process. Due to the fact that they are very private and private, especially within the EU, forecasting models cannot be optimized and used to their full potential. A particular contribution to their improvement would be a study that analyses a set of data combined from several modules of telecom equipment, e.g. a combination of traffic data, location and movement data (GPS coordinates), mobile wallet transaction data, applications used, mail client, etc. Given the sensitivity of such data, it needs to be sufficiently protected and anonymized before being subjected to processing. Finding a solution for accessing and processing such information would have a huge impact on the concept of modern digital marketing: the formation of a single customer profile, in which all the actions of the individual customer are traceable across different platforms, enriched with the power of analytics and predictive algorithms, would make it possible to absolutely personalize business and public activity down to the lowest level of the individual.

IV. Publications related to the topic of the dissertation thesis

Scientific papers:

Trichkov, D., Understanding and analyzing the Sub Sahara Africa market potential for digital financial services, Annual research almanac of doctoral students, 2018, Book 14, Academic publishing house “Tsenov” – Svishtov, ISSN 1313-6542

https://almanahnid.uni-svishtov.bg/title.asp?title=1395

Trichkov, D., Bulgarian telecoms market challenges and opportunities, Annual research almanac of doctoral students, 2019, Book 15, Academic publishing house “Tsenov” – Svishtov, ISSN 1313-6542

https://almanahnid.uni-svishtov.bg/title.asp?title=1533

Scientific studies:

Trichkov, D., Customer demographic segmentation based on telecom behavioral data, 2020, Book 16 , Academic publishing house “Tsenov” – Svishtov, ISSN 1313-6542

https://almanahnid.uni-svishtov.bg/title.asp?title=2644

IV. Statement of contributions to the thesis

First: Proposed new concept of cross-industry business models implemented based on alternative use of big data is proposed.

Second: Shown relationship between mobile telecommunication service consumption and individual customers' reported income by demographic cross-section is empirically established.

Third: An author's model is proposed for the process of operation and allocation of responsibilities in a financial services platform based on the combined information resources from telecommunication and banking institutions.

V. Statement of originality

(According to Art. 68, Para. 2 of the Regulations for Academic Staff Development at D. A. Tsenov Academy of Economics)

The 218-page dissertation entitled "Product Diversification through Big Data Analysis" is authentic and represents the author's own scientific production. It uses the author's own ideas, texts and visualization through graphs, charts, tables and formulas, and complies with all the requirements of the Copyright and Related Rights Act by properly citing and referencing other authors' thought as well as data, including:

1. The results achieved and contributions made in this dissertation are original and have not been borrowed from research and publications in which the author is not involved.
2. The information presented by the author in the form of copies of documents and publications, personally compiled references, etc. corresponds to the objective truth.
3. The scientific results obtained, described and/or published by other authors are duly and extensively cited in the bibliography.

This is to certify that to the best of my knowledge; the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Date: 01.12.2021 PhD Candidate: ……

(Dimitar Trichkov)