Performance of Crowdfunding in Agricultural Campaigns

MSc Research Project
Msc Fintech

Ogochukwu Pamela Ezeonugo
Student ID: X18211534

School of Computing
National College of Ireland

Supervisor: Noel Cosgrave

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Ogochukwu Pamela Ezeonugo ……. ………………………………………………………………………………………………………
Student ID:	X18211534 ……………………………………………………………………………………………………….……
Programme:	MSc fintech …………………………………………………………… **Year:** 2020…………………………..
Module:	Research Project ……………………………………………………………………………………………………….………
Supervisor:	Noel Cosgrave ……………………………………………………………………………………………………….………
Submission Due Date:	15-08-2020 ……………………………………………………………………………………………………….………
Project Title:	Performance of Crowdfunding in Agricultural Campaigns ……………………………………………………………………………………………………….………
Word Count:	8282 ……………………………………… **Page Count**……………22………………………………

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ………

Date: ………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Performance of Crowdfunding in Agricultural Campaigns

Ogochukwu Pamela Ezeonugo
X18211534

Abstract

In the sphere of agricultural sector financing, the introduction of crowdfunding has served as a breakthrough in helping the financial deficiency in this sector. This study utilized previous works to understand the factors that influence crowdfunding success in agriculture campaigns. Using samples of agricultural campaigns launched on the Kickstarter platform, under the food category, the study uses regression analysis to model and investigate the factors that affect both the success and success rate of a campaign. The results of the study indicated crowdfunding performance in agricultural campaigns is affected by specific factors which project creators fail to handle when seeking for funds on a platform. The findings suggest project creators may need to use different strategies in achieving successful projects. A relatively achievable funding target and social circle expansion both offline and online to increase the number of contributors needs to be put in place at the onset.

Keywords: Crowdfunding, campaign, projects, Kickstarter, platforms, backers, agriculture

1 Introduction

The growth of the financial industry in recent times and the evolving digital market has created a change in the industrial cycle. This change has brought about new financing surveys such as crowdfunding. Crowdfunding is the use of small amounts of capital from many individuals to finance a new business venture. The transaction value of crowdfunding is projected to reach US$940.9[1]m in 2020 alone. Crowdfunding has created a new financing model for individuals with limited access to financial support and resources to be able to obtain financial assistance from various individuals through online platforms (Mollick,

[1] https://www.statista.com/outlook/335/100/crowdfunding/worldwide

2014). The amount of money crowdfunded globally increased from $16.2billion to 34.4 billion. This is an increase of 112% from 2014 to 2015[2]. These figures prove the viability of crowdfunding as a valuable resource tool. Crowdfunding has been used in sectors such as real estate, agriculture, sports, entertainment, art, and healthcare.

On agriculture, this report aims to examine the various ways crowdfunding has performed in the agricultural sector over the years. Agriculture is a vital part of human history. There has and will always be a need to provide food for human sustainability. Today, the agri-food sector has become an important key player in the role of economic growth in most developing countries and serving as a booster for their GDPs (Chang, 2018; Thaker et al., 2020). Small scale agricultural ventures have little access to capital when compared to large-scale agricultural ventures. Crowdfunding(CF) can prove to be another source of funds for these small-scale ventures.

The agricultural sector is known to be a capital-intensive industry due to the consistent operational costs needed to keep industries and farms running. The use of new technology to increase efficiency and produce more food is not an easy task. This requires a constant flow of cash which Agrarian entrepreneurs and small-scale farmers struggle with. Government grants and loans from financial institutions have proved to be non-feasible due to their low volumes and inconsistency. While various investors are willing to provide such capital for the agrarian entrepreneurs, these investors assume the payback period in the agri-food sector is longer and less likely to pay off and the nature of the sector is risky due to weather conditions (Mardhiyyah et al., 2020).

This is where crowdfunding has come in. Soliciting for funds and financial assistance from the public seems to be an easier and less cumbersome method to raise funds via the Internet, Giving agrarian entrepreneurs and small-scale farmers' access to a sophisticated toolbox for finding new ideas, testing new products, studying new needs niches, and creating customer loyalty (Misso and Cesarrati, 2017).

This report looks at the various models through which agrarian entrepreneurs and small-scale farmers can properly crowdfund to reach their goals. Examining 4 different models, this report looks at the strengths and weaknesses of these models, thereby explaining why there is a low performance of crowdfunds in the agricultural sector. Although there has being a low success on crowdfunding in agriculture this has led to the creation of agricultural centred

[2] http://reports.crowdsourcing.org/index.php?route=product/product&product_id=54

platforms such as AgFunder, Barnraiser and Woopfood in the United States and Europe. AgFunder has raised over $9.3 billion over the years in agricultural projects[3].

The current challenging issue is most crowdfunding platforms(CFP) make use of funding models which are not favourable to agrarian entrepreneurs (Pronti and Pagliarino, 2018). Nowadays, over there has been an increase in the number of entrepreneurs seeking investors through crowdfunding, so improving current models on crowdfunding success has become very important.

1.1 Motivation

- The need to add to existing research work in this space.
- To address the current state of deficiency being experienced in accessing funds for facilitating agricultural projects.
- The need for increased awareness on online sourcing of funds for financing agricultural projects via crowdfunding markets/platforms.
- The need to improve the general understanding of crowdfunding in the agricultural sector and its operational activities.
- To provide a cheaper source of funding for agricultural entrepreneurs.

1.2 Research Objective

To predict the success of agricultural crowdfunding projects using text analysis and machine learning techniques.

1.3 Research Question

What are the key determinants for successful crowdfunding campaigns in the agricultural industry?

1.3 Contribution

The study gives a general insight into the existing literature on crowdfunding in agriculture. Although empirical literature is still developing the study, the study can create awareness for agarin entrepreneurs in addressing their funding needs and providing them with a guide on how to maximize their chance and degree of success.

1.4 Limitation of the Study

The current literature on crowdfunding success has widely ignored agricultural projects in their works, which usually include technology, art and fashion. A few studies have provided an insight into the factors that contribute to crowdfunding success in agro projects. However,

[3] https://agfunder.com/

empirical research in the context of agriculture is extremely limited and to the knowledge of this study and with a large enough sample to reach a generalizable conclusion.

1.5 Structure of the Research Study

The paper is divided into the following sections the related works which provide academic literature on crowdfunding; the methodology sections provide details of data insight, techniques and evaluation metrics that would be used in carrying out the study; the design and implementation section gives a detailed report on every procedure taken to handle the data before being modelled. Lastly, overall results report which will be used in the discussion and conclusion of the study

2 Related Work

The birth of crowdfunding has led to other aspects such as crowdsourcing, the crowd-voting, crowd-creation, which have one way or the additional benefit from the crowd (Crosetto and Regner, 2014). Crowdfunding is a process in which capital seekers either as individuals or an organization request for financial assistance from a group of persons who is anonymous through a digital platform with the absence of financial intermediaries (Schwienbacher and Larralde, 2010). Belleflame et al., (2014) and Mollick (2014) further state that crowdfunding generates the same amount of financial resources as the funds obtained from a bank and interest in what is created will motivate people to invest or contribute to a crowdfunded project. The ever-growing use of digital technology has led to the creation of a digital market place which serves as an enabler to the processes in which crowdfunding activities and key players have made efficient use of since the emergence of this new funding model (Anshari et al., 2019).

Donation-based Model-The donation-based crowdfunding is one of the most effective models in the platform. It involves requesting for small amounts of money from many individuals in a crowd with the little bits leading to high volumes of income. In the agricultural sector, Mollick (2014) indicates that donors use the model to assist communities in the funding process. The online platform has developed an excellent opportunity for financial acquisitions.

Reward-based Model: The reward-based crowdfunding is useful for start-ups, (Agrawal, Catalini & Goldfarb, 2014) imposes that projects based on agri-food practices are often adequately funded in this method. In the process, donors acquire awards from the amounts they donate. The rewards are usually adopted as appreciation benefits for the services delivered to those in need. Instances of rewards are also meant to entice donors to provide valuable resources for the project (Hemer, 2011).

Equity-based Model: Research by Gerber, Hui, & Kuo (2012) presents equity crowdfunding model as majorly meant for the small and medium enterprises in the situations they require vast amounts of funds to begin large-scale operation. Small-scale farmers are usually the beneficiary of such funding; especially in the instances, they have potential land for further exploration. In this case, the donors acquire part of the ownership of the project. The percentage earned usually differs based on the nature of the investment (Ahlers et al. 2015). This takes a similar process as the ordinary shareholding in the stock market. However, this process usually takes place in the unlisted company form. The risk associated with the model is generally realized when the donor's rights are transferred to another investor where the amount of the share is reduced.

Debt-Based Model: The debt crowdfunding model involves collecting money with the promise to refund the owners at an expected period (Schwienbacher, & Larralde, 2010). This model is usually preferred by investors who avoid equity and sharing their operations. Donors often require the specific date for repayment of debts as they take the same form as the ordinary loans, although diversified to some extent. In the agri-food business, (Kuppuswamy, & Bayus, 2018)explained that the stakeholders may receive the funds and propose to repay after the production process is over.

Funding dynamics is a key aspect in crowdfunding models and involves a mix of two which are *keep-it-all* and *all-or-nothing* funding models. The use of this funding dynamic is what determines the structure of every crowdfunding platform. The *keep-it -all* funding dynamic is commonly used by platforms which are donation-based inclined and grants project creators the right to access the funds contributed whether project goal is achieved within the targeted time frame (Cumming, Leboeuf & Schwienbacher, 2015). *All-or-nothing* funding dynamic is usually adopted by platforms that are equity and reward-based, the funds can only be accessed by the project creator on the condition that targeted funds is achieved before the deadline of the campaign; failure to meet the target by the project creator the funds will be returned to the backers. In recent times, some platforms offer a mix of the two to reduce the rate of unsuccessful campaigns by project creators (Filmonova et al., 2018).

2.1 Crowdfunding Participants

Crowdfunding has been regarded as one of the efficient techniques for enhancing cash flows for a potentially developing enterprise but facing hardship in other available means

(Hildebrand, Puri, & Rocholl, 2017). When it comes to crowdfunding, there are two key players; Capital seekers and Capital providers. Both are equally important because they require each other to coexist. Capital seekers are the project creators who seek to select the preferred funders and the type of relation which adequately solves the problems presented. Capital providers are those who fund these projects and are offered various choices to fit into the provided models, which include the possibility of adopting a loaning method type of funds and grants which are free. The agri-food sector is one of the major projects the funders are willing to support for the value provided to the society. Besides the traditional practices for offering support, capital providers efficiently provide their services through the internet, thereby quickly pooling their resources for the benefits of those who need assistance (Ellman, & Hurkens, 2019).

In the agriculture and food industry, investors are prone to using technology in their activities and connections are essential in the online monetary acquirement. In the aspect of providing support for the agri-food projects, undertakings become one of the best ways the funders engage with the project owners (Hornuf, & Schwienbacher, 2017). Crowdfunding is one of the riskiest practices, especially where the funders adopt the equity and debt alternatives, as they are always not sure of the returns. They face the risk of being transferred to other project developers leading to the drop of share capital. According to (Pronti and Pagliarino, 2018) the agri-food industry could be one of the risky areas of investing through the debt and equity funding processes because of the unpredictability and changes in seasons for food production. The recent events are clear examples to learn from based on changing weather patterns, natural calamities, and unpredictable seasons (Pacchi, & Pais, 2020).

2.2 Determinants of Success in Crowdfunding

It is not always that the project developers will always gain the support required from the funders. As in any other market platforms, crowdfunding requires that the bidders provide the most pleasing information capable of luring the interest of the funders (Cillo, Cardinali, & Bertoldi, 2017). The project and creators are always the critical determinants of the awarding decision in the crowdfunding process. The reliability of the information provided to the funders should have a verification method where they can depict the awarding decision or forego the procedure. Carolan (2017) indicates that issues of trust and honesty are usually necessary for successful interaction between the two parties. The funder may require backup evidence from the authorities to clear verify the details. If the critical information is included, the next step would be requesting for pictorial presentations that show the validity of the indicated project proposal (Anshari et al. 2019). This is key to the reduction of the information gap across the developers and funders. Ward and Ramachandran (2010) explained out that the use of peer effect by backers in making funding decisions which occur by looking at other actions of other backers towards a campaign; supporting this notion (Yum et al., 2012) using binary logistic regression in a study termed it the herding behaviour in crowdfunding; whisch is a given situation were backers pick interest in campaigns that have high funding pledges. Besides all the material evidence (Levy, Navereau, & Triboulet, 2018) the language used in the request platform is another crucial area of concern. The use of selective words that exhibit persuasiveness easily depicts a favourable judgment and the decision to fund the project (Mariani et al., 2016). Wang et al, (2017) in a study using text mining approach point out that proper description of a project through the blurb would likely have an impact in attracting more capital providers but even with a high positive sentiment towards a title of projects this would not always be the case. Mitra and Gilbert, (2014) carrying out an analysis of the number of words and phrases suggested that the use of more persuasive phrases can attract more project backers. In progressing projects, highlighting the various achievements motivates the funders as they gain the willingness to take part in such a development. This is usually meant to ensure that no resources are wasted from unpredictable projects (Mitra, 2012). For instance, the agricultural project represents one of the most illustrative sectors to provide the progress achieved so far. Funders in crowdfunding are also influenced by relating the indicated project to others in terms of failure or success. They rely on their peer action in making the decision. The information may also be used in determining how a project may or may not work and decide on the best route to take (Cumming, Leboeuf, & Schwienbacher, 2015). Achievable expectations from the project developers are usually an important area that determines the funders' decision. Instances when the demands can be met, leading to possible consideration for awards. Beaulieu, Sarker & Sarker, (2015) indicate that cases of rewards are likely to result in successful funding when the funders prospect the benefits accrued to them when striking the deal. Although the sharing of information is important in the success of crowdfunding (Xu et al., 2014) using statistical analysis indicated that the use of high visual representation, a constant update which can both be offline and online is paramount in the success of a project getting funded. (Hui et al., 2014) conducting a qualitative analysis found out that most project creators that had successful campaigns made an effort to have a relationship or update projects backers both offline and online.

Filimonova et al., (2019) using the application of correlation -models on two crowdfunding platforms in Russia points out that agrarian entrepreneurs need an alternative source of finance which could be crowdfunding. Also, the study further asserts that agrarian entrepreneurs lack a basic understanding of crowdfunding mechanism and thus most projects end up not being funded, this was the idea was further supported by Chang (2018) stating agricultural projects on crowdfunding platforms lack distinctive features from other projects to attract sponsors. The capita income of the population and information-communication is a central key to the success of crowdfunding in any geographical location (Agrawal et al., 2011; Filimonova et al., 2019).

Chang (2018) using binary logistic regression on crowdfunding projects in the agro-sectors in Asia argues that for crowdfunding to be successful in agricultural projects, project backers must diversify agro-projects into other categories to attain project goal. On the contrary (Misso and Cesaretti, 2017) argues that even with the diversification of projects into different categories crowdfunding platforms still play a key role in selecting whether a particular project is worthy of financing. Filmononva et al &Ng et al., (2018) indicated that project backers fail to determine the perfect crowdfunding strategy for agricultural projects to be successfully funded. This strategy should include social impact, project scale and a vivid description of project viability.

Thaker et al. (2020) created a model and suggested for crowdfunding to be successful in agricultural projects, crowdfunding platforms must be infused with Islamic banking which eliminates the payments of interests, this argument is supported by (Pronti and Pagliarino, 2018) that points agrarian entrepreneurs are usually faced with the challenge of the model used by the crowdfunding platform. Confidence for funding increases in situations where successful projections are made (Xu et al. 2016). Project developers have also gained successful funding when they undertake promotion, primarily through social media platform where the information reaches as many people as possible.

In general crowdfunding success in agricultural campaigns is within the control of each key player and how they handle the different factors that contribute to the successful attainment.

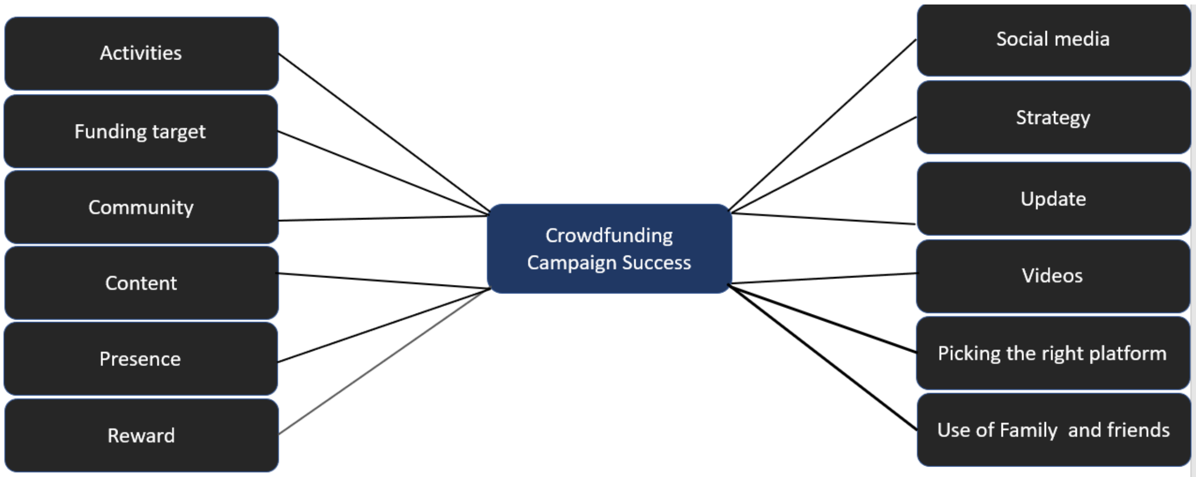


Figure 1: Influence of Crowdfunding Success

2.3 Kickstarter

Kickstarter is a reward-based crowdfunding platform that mainly focuses on creative projects since the creation of the platform in 2009; it has provided funding for different projects such as Pebble. The platform supports creative projects which vary from Music, Arts, Fashion, Games, Film, Comic, Dance, Video, Food and Publishing. The platform has funded over 183,978 projects and rose over $5 billion as of June [4]2020 for projects on the platform. The platform operates on an *all-or-nothing* model to fund projects. In participating in Kickstarter as a project creator individual can register on the platform without payment of any fee. The platform allows individuals to either be project creator or project backers. The maximum pledge on a project is $1000[5] and a project on the platform runs for about 1 to 60 days but the platform usually gives a recommended time frame of 30days. The platform has no geographic restriction on project backers, but project creators must be from specific countries (Kuppuswamy and Bayus, 2013; Marom et al., 2014).

[4] https://www.kickstarter.com/
[5] https://help.kickstarter.com/hc/en-us/articles/115005066393-How-do-I-pledge-

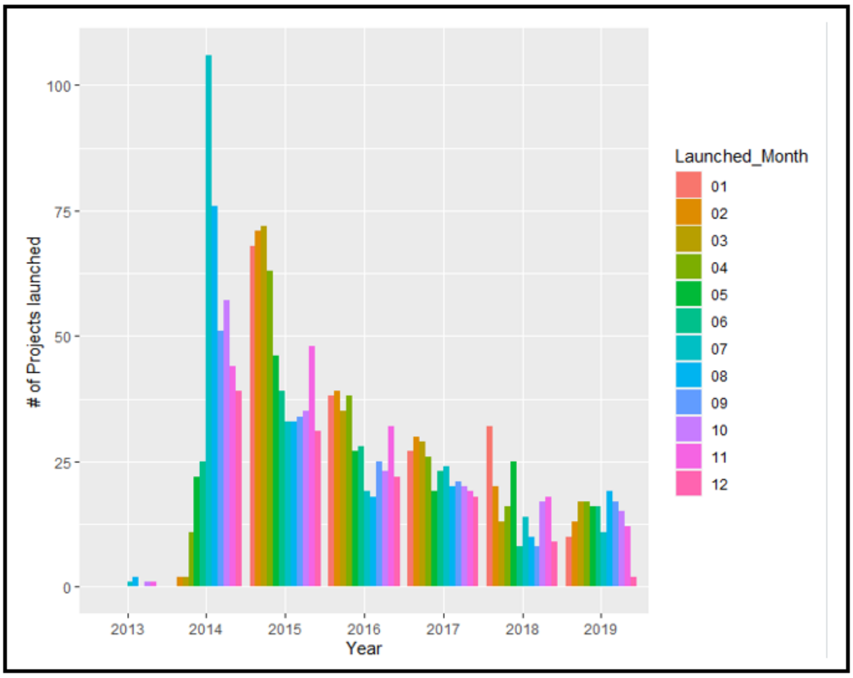


Fig 2: Agricultural Projects on Kickstarter platform

3 Research Methodology

This section provides a detailed description of the data and methods utilized in the study and motivations for its selection. In carrying out an empirical analysis in this research the study follows the Cross-Industry Standard Process for Data Mining (CRISP-DM).

3.1 Data Understanding

The dataset acquired for the study was extracted from a WebCrawler that hosts Kickstarter data. The data set selected was from June 2013-January 2019, containing 37 columns and 18669 projects. The data contains different project category that has a different subcategory. The category selected for this study is the project category "Food" which has several subcategories. This study will focus on the subcategories of "Farm" and "Farmers Market" this can be justified in the works of (Chang, 2018; Filimonova et al., 2019) in conducting empirical analysis in the study.

3.2 Variable Review

i) Dependent Variables

State: The "**State"** which is a continuous variable that indicates whether a project reached or exceeded the funding target (1=" successful", 0=" failed").

Success rate: The success rate of agricultural projects in crowdfunding will be measured with the two variables funding pledged divided by funding goal (Calic and Mosakowski, 2016). Every project campaign can either exceed its funding target or reach the specific target.

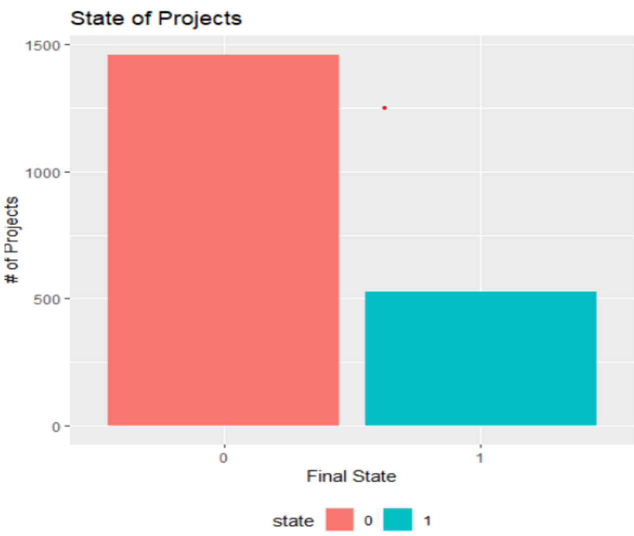


Fig 3: State of Crowdfunding Projects

ii) **Independent Variables**: The justification and selection of independent variables are based on previous academic and empirical research in prediction crowdfunding success. Tu et al., (2018) classified theses variable using hypothetical impact analysis.

Backers Count: The total number of capital providers for a project on the crowdfunding platform. Projects are likely to succeed if the number of backers is high (Ahlers et al., 2015).

Goal: This is the target amount the project creator seeks to raise to start the project peached on the platform. The lower the goal of the project the easier it is for the project to reach the funding target (Ahlers et al., 2015).

Blurb and name length: This usually describe the project by using words that can attract backers.

Duration: This is usually the time from when the project is peached on the platform till the expiry date. Mollick (2014) suggested that the duration of a project is not a guarantee of success.

Staff Pick: The staff in a crowdfunding organization can be bias towards projects and so can ensure the projects are viewed by multiple backers visiting the platform**.** (Lelo de Larrea et al., 2019)

3.3 Data Preparation

The data was extracted from compressed file consisting of 45-56 Microsoft excel files with crowdfunding projects information. The data was sampled selecting information strictly related to the category of farmers, farm and farmers market. The data is retrieved from a

WebCrawler and cleaning would be done using the R Studio and metadata was non-existent. The cleaning of the collection will be discussed in detail in section 4.

3.4 Modelling

This section includes the organization and execution of the required models in this research. Logistic Regression, Support vector machine (SVM), Multiple Linear Regression and Sentiment analysis are used to predict the success of crowdfunding in agro projects. The models are tuned to achieve optimized accuracy.

i) Logistic Regression: This machine learning technique describes data and explains the relationship *between a "dependent binary variable and other independent variables*". This modelling technique is not particularly considered as a regression method even though its name suggests so. It is, however, a classification technique that has a probabilistic output that an input belongs to a specific class. Logistic regression is a "*non-linear transformation of the logistic regression model*" which has an S-shaped logistic distribution function with "*predicted probabilities between 0 and 1*" (Markas, Wang, and Tseng, 2019). Furthermore, as explained in Korkmaz, Guney and Yigiter (2012), this method is key in "*categorization and process of appointment*"; it is through estimation of dependent variables that categorization is attained in line with the probability rules. The interpretation of the results obtained from this method is simple and easy.

ii) Multiple Linear Regression – This is a statistical method used in estimating variables relationship. It provides an analysis of how the dependent variable and one of the independent variables and from this the linear relation equation is formulated. However, when in solving regression problems, if it involves one dependent variable and more than one independent variable, it is referred to as multilinear regression. Regression is used as a model when determining the correlation between variables which have "*cause-effect relations*" (Uyanik and Guler, 2013). Furthermore, Marill (2004) explains that this model is developed based on simple linear regression. It is a" *generalization of simple linear regression in which there is more than one predictor variable*". Some of the advantages of this model are that it has a highly accurate and provides a clear understanding of the relationship between the variables.

iii) Sentiment Analysis: Opinion mining is one of the processes of using algorithms to organize and categorize text, sentences, reviews or books. The objective of carrying out sentiment analysis is to determine the opinion of an individual towards a specific product, service or feature (Luo et al., 2013). Farhadloo and Rolland (2016) explain one of the advantages of sentiment analysis it can serve as a complement to other systems from question answering to a recommendation. Sentiment analysis can be carried out using two basic approaches the lexical method which is an approach that classifies the text using a dictionary or words and breaking down the text into positive and negative or classifying words into positive and negative (Wang, Zhu, Wang and Wu, 2017). The other approach is the use of a machine-learning algorithm to break down the text into different parts of speech and create a vector for a popular word(Bhadane et al., 2015).

iv)Support Vector Machine: SVM is machine learning techniques that handle both classification and regression, and efficient in handling predictive performance. The SVM works by searching for the optimal surface with the use of algorithms and kernels. The advantages of using the SVM is the use of its sparse technique that allows it to carry out a non-parametric method on the training data (Awad and Khanna, 2015).

3.5 Evaluation

After the implementation of the models, the efficiency and performance of each model would be evaluated using the following metrics: Adjusted- R^2, F-statistics, accuracy and confusion matrix. Visualization of sentiment will be done using a word cloud. Error value with lower figures signifies the performance of the model a high adjusted R-squared explains there is a correlation between the observed and predicted values if the values are between 0.7 and 1.

4 Design Specification

The data is sourced from Webrobots[6] and it covers agricultural projects on the crowdfunding platform for the period 2014-2020. The data is cleaned/transformed; sentiment analysis is done via Bing, NRC and Afinn. Thereafter, the data is ready for regularization and variable selection using the Elastic Net technique. For this thesis, cross-validation and confusion matrix is used in evaluating the model. The models used in this analysis are Logistic Regression after which the results will be presented.

[6] https://webrobots.io/kickstarter-datasets/

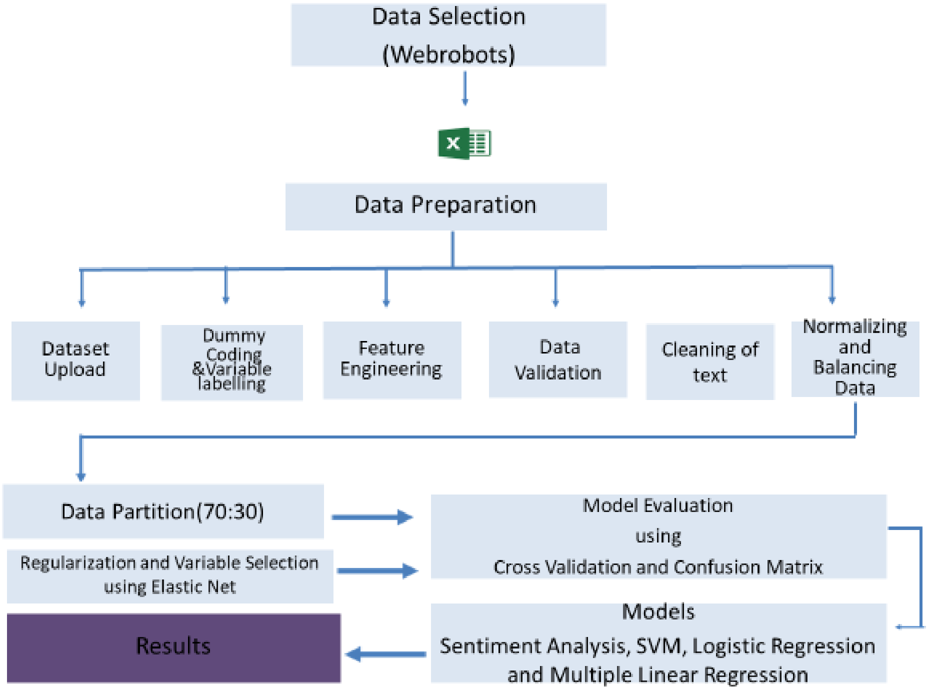


Fig4: Design Architecture.

5 Implementation

5.1 Data Selection

The data set used for this study was retrieved from[7], sourcing of the data was done from various files selecting specifically two categories which were farmers market and farm only, that were needed in the study. The total numbers of rows in data set were 18,669 and containing 37 columns of which some of this column contained irrelevant information. Compilation of the data was done on the Microsoft excel before the data was imported into the R Studio for data cleaning processes. The data set had no metadata upon retrieval; therefore, metadata was created.

[7] https://webrobots.io/kickstarter-datasets/

Field	Data Type	Description
ID	character	ID of Kickstarter Project
Blurb	character	Description of project
duration	numeric	Active Duration of the Kickstarter Project
goal	numeric	Goal of Kickstarter Project in original currency
launched year	character	Year when Kickstarter Project was launched
Launched month	factor	Month when Kickstarter Project was launched
deadline year	factor	Year when Kickstarter Project ended
deadline month	factor	Month when Kickstarter Project ended
currency	factor	Currency of amount pledged
pledged	numeric	Total amount pledged in original currency
category	factor	Category of Kickstarter Project
name	character	Name of Kickstarter Project
backers	integer	Number of Backers
state	factor	Final State of Kickstarter Project (success or failed)
country	factor	Country where Kickstarter Project was launched

Table 1: Dataset Description

5.2 Data Pre-processing and Transformation

In carrying out any machine learning analysis it is essential to process, clean and transform the data before fitting the model. The data used in the study consist of both categorical and numerical features. Independent categorical features such as “country” and “currency” had multiple levels, and this was broken down into five-factor levels by selecting major countries and currencies. The category countries were releveled as “US”,” GB”,” CAD”, “EUR” and “Others”. Also, category currency was releveled into five factors as “currencyAUD”.” currencyUSD”,” currencyGBP”,” currencyEUR” and lastly “currencyOthers”. Categorical features of both independent and dependent variables were dummy encoded assigning using a matrix function to create these levels to new data frames assigning “1” or “0” to “TRUE” and “FALSE”. The new data frames created containing the dummy encoded variables were now combined to the original dataset this led to the increase in columns and variables. Features such as “launched_at”, “created_at” that were represented as strings or timestamps in the data were converted from strings to date and broken down into months, days and years this also led to the introduction of more variables and columns. Since the data had categorical features the next step in our exploration was assigning labels to specific features in the data set such as “date”, “backers count”,” id” and “state” which was assigned a binary label “0” or “1” to “failed” and “successful” for dependent variables. Missing values in the dataset were 9945 which was from columns provided in the data that had no information upon retrieval these columns which are “permissions”, is_ starred, is backed and “X” are all denoted with multiple “NA”. Therefore, no computational analysis was carried out to handle missing data and these columns are eliminated from the data set. In handling duplicates this was done using the “Id” column for each project this is done to avoid a mixture of different figures that signify different value measurement. The uniqueness of a project is differentiated by the projects ID number on the Kickstarter platform.

Feature engineering is implemented on the independent features that were specifically needed in carrying out the modelling technique and have been used in previous studies such as “duration” was created using the variable” launched_at and “deadline”, while “name length” was created from the variable “name”. The success rate is another variable that was created using funding pledged divided by the finding goal. Visualization of the data is done with the use of a graph to show w various features of the data from dates of the project and funding patterns. Also, a new dataset was created from the original dataset to investigate the distribution of the dataset. A correlation analysis is carried out to determine the variable in the data that are highly correlated and might create issues in our modelling process, Fig 4 below shows the correlated variables**.** The removal of irrelevant columns containing URLs and repetitive information from other columns, but which were described using a different name in the data was removed to avoid the issue of multicollinearity or noise among the features. After all these processes the data was re-ordered and scaled to ensure all numeric features were within the same threshold. The next stage multiple datasets were created to test the model assumptions, sentiment analysis and modelling.

Variable	N	Mean	Std.dev	Sum	Median	Min	Max
Backers count	1988	46.74	263.1377	92926	5	0	10308
Goal($000)	1988	88.755	694838.2	17644.585	1.05	1	2.2
Success rate(%)	1988	0.335444	0.52456	14075.989	0.0191149	0	1.0
Duration	1988	30.76	206.41	61156	30	-1568	1621
Name length	1988	5.574	2.4997	11082	5	1	1.4

Table 2: Descriptive Statistics

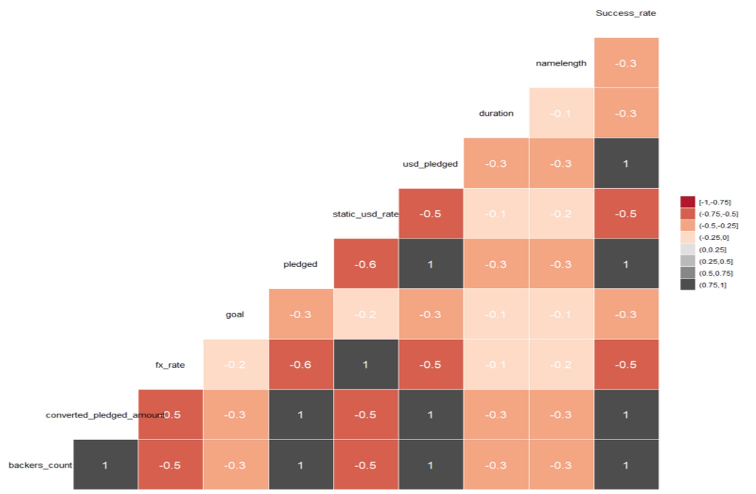


Fig5: Pearson Correlation Plot

5.3 Model Diagnostics

In carrying out any given regression model there are specific assumptions to be tested to ensure the model results produced are unbiased.

i) Influential Values: Using the cook's Distance to check the influence of an outlier, using the two-thumb rule (Dhakal, 2017),

ii)Test for Multicollinearity: This was done using the Variance inflation factor which is used to measure multicollinearity among variables(Hsieh et al., 2003).

iii)Test for autocorrelation: This done with the Durbin- Watson test for autocorrelation used to verify if there is a serial correlation among variables[8].

5.4 Cleaning Text

The creation of a new dataset was done to specifically handle text mining analysis in the model. In the processing of text, this study uses two approaches which are tokenization and removal of stop words. Tokenization is applied to one of the features of the data set "Blurb" and this is done to separate words from another and represented as a token. This was done using the function in the R library (tidytext). The other steps are the removal of stop words is done to eliminate common terminologies used in the English language. The three lexical approaches were used in analyzing the text

5.5 Data Sampling and Partitioning

In carrying out any modelling analysis, the data set is partitioned into a training and testing set, thus the research uses a ratio of 70:30 for the data partition. The issue of class imbalance

[8] https://www.statisticshowto.com/durbin-watson-test-coefficient/

is specifically common to categorical dependent features and this dataset there is a huge difference between successful and failed projects. To address this issue of class imbalance data sampling must be carried out. There are two methods of handling this issue which could be either by oversampling or under-sampling of the class and this would be done in the RStudio. In tackling this issue for this dataset, the SMOTE function in R is applied and tuned with the right parameters creates an equal balance between the two classes by using both oversampling and under-sampling techniques**.**

5.6 Variable Selection and Regularization

Variable selection and regularization are important in carrying out a regression analysis, this done to eliminate variables that tend to create noise in the model. This study uses th**e Elastic Net Regression** technique to regularize the data set and for variable selection. The elastic net is a hybrid regression method which is a combination of the lasso and ridge regression; it is usually applied as a regularization and variable selection technique (Chen et al., 2019). One of the advantages of using the elastic net in applying shrinkage is the ENET handles data with high dimensionality, thereby correcting the drawbacks of the lasso regression (Doreswamy and Vastrad, 2013). In regularization of data the ENET uses to penalties the lasso (ℓ_1) and ridge (ℓ_2), the elastic net is a simple ridge regression when α=0 and a lasso regression when α =1. The use of the (ℓ_1) and (ℓ_2) for variable selection and random sampling is an advantage that makes the ENET efficient in regularization (Ogutu et al., 2012). The lambda is a determinant of the overall strength of the penalty while the selected α is a high-level parameter which can be applied in testing of a model. The highest fraction deviance explained is between 2 and 12 for a non-zero variable with the highest level of alpha at 0 and lambda at 0.2. The variable which was reduced to zero by the elastic model is eliminated and the dataset is partitioned using a 70:30 split again to carry out the modelling techniques. Each modelling techniques is carried through cross-validation and confusion matrix. The multiple linear regression model is evaluated through the coefficients and R-squared.

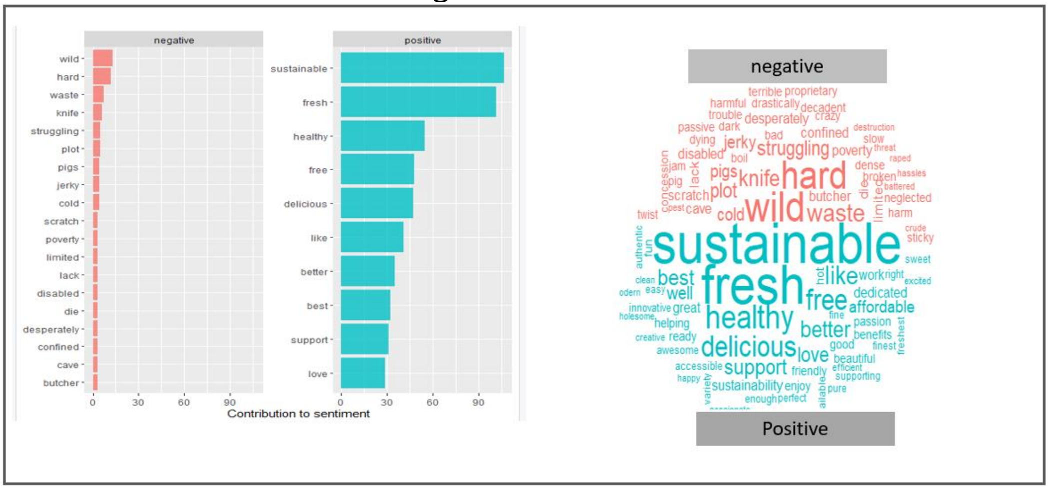
6 Evaluation

The result of this study comprises of three sets a sentiment analysis using the lexical approaches to classify people’s opinion into negative and positive towards crowdfunding projects. The use of logistic regression and SVM to classify projects into success and failure and the other is the use of multiple and linear regression to investigate probability.

6.1 Sentiment Analysis

These are words using the lexical based approach that potential backers on a CF platform find as positive and negative words that will affect the funding of the project. From the figure below the word, “sustainable” is the most common word in the description of an agricultural campaign on a platform.

Fig6: Word cloud



6.2 Predictive Analysis

The table below represents the accuracy result from the tuned model and the confusion matrix. Both modelling techniques were tuned using different parameters in carrying out the prediction.

Model	Accuracy	AUC	Kappa
SVM	0.8569	0.8225	0.7139
Logistic regression	0.8744	0.8408	0.7488

Table 4:Performance Matrix

Model	SVM Actual Failed	SVM Actual Succesful	Logistic Actual Failed	Logistic Actual Successful
Predicted :Failed	421	43	416	36
Predicted:Success	28	104	33	11

Table 5: Confusion Matrix

6.3 The determinant of Funding Success in Agricultural Projects

Table 4 represents the model with the use of binary logistic regression, after testing multiple variables in the model and eliminating all the factors having sig < 0.05. The model finally determined three factors: goal, pledged, name length, backers count, other currency and country US. The model using the pseudo-R^2 explains 63.4% of the variation of funding success. The project with long name length would decrease the likelihood of funding success. Projects with an enormous number of backers would be successful in projects.

Table 4: Logistic Regression

Variables	Coefficients Estimate	Coefficients Std-Error	Coefficients Z-Value	Coefficients Probablity
Intercept	- -0.7360	0.1937	-3.799	0.000145 ***
Backers count	-571.2660	53.1151	-10.755	< 2e-16 ***
goal	633.3468	29.8096	21.246	< 2e-16 ***
Name length	-0.8753	0.3407	-2.569	0.010202 *
Currency Other	1.4638	0.5324	2.749	0.005969 **
Country Us	-0.3092	0.1683	-1.837	0.066145 .

Significance codes: ***0.001, **0.01, *0.05, 0.1

Table 5: Multiple Linear Regression

Variables	Coefficients Estimate	Coefficients Std-Error	Coefficients Z-Value	Coefficients Probablity
Intercept	0.001790	0.001100	1.627	0.1039 *
spotlight	0.008886	0.001002	8.864	< 2e-16 ***
Staff pick	0.007080	0.001685	4.201	2.76e-05 ***
Name length	-0.005326	0.002648	-2.011	0.0444 *

Table 6 above the modelling is performed using success rate as the dependent variable, testing of different hypotheses is done and to determine factors that have a significant value of less than $p< 0.05$. The model was statistically significant at an (Adjusted $R^2=0.797$, F-statistics=41.94 and $p<0.001$).

$$\text{Success rate} = b0 + b1*\text{staff pick} + b2*\text{spotlight} + b3*\text{name length} \qquad (1)$$

The model assumptions the result of regression analysis, the model assumptions were met. VIF values were below the threshold of 5, which so the assumption of multicollinearity has been met. Durbin-Watson statistics (1.705404) fell within an expected range, thus indicating that the assumption of no autocorrelation of residuals has been met as well. Using Cook's distance to for influential cases, the result had a maximum value of 0.94, so the assumption of influential values was met, and the data is normally distributed.

6.4 Discussion

This research summarizes the factors affecting crowdfunding campaigns in the agri-food sector by looking at the players of a CFP. The results indicated that there is a ***significant relationship*** between the state of a crowdfunding campaign and the number of backers. This is in line with previous studies by (Chang, 2018) that for the status of the agricultural campaign to change from failure to success, the number of backers must increase. The higher the number of backers the likelihood a project would get funded within the expected period. The previous studies by (Tu et al.,2018) demonstrated that a campaign with a minimum number of backers increases the success rate due to large contributions compared to a campaign with more backers and less contribution.

Secondly, the length of the description of a campaign has a ***significant relationship*** with the state of a campaign. Although Koch & Siering (2015) explained that this usually has a positive impact on the success of a campaign, this fails to hold in terms of an agricultural campaign. From a good sample of results reviewed, it was observed that people are more likely to donate if they see words in the description of projects that sound positive especially if these words speak along the lines of better environmental sustenance An agricultural campaign needs a potentially large amount of capital providers to be able to meet the funding target and the funding goal is a key factor. The higher the funding target, the less likely a campaign would be successfully funded. Projects with low funding targets are likely to get funding sooner than projects with a higher threshold because backers are more likely to get attracted to lower targets. This is in line with research by (Tu et., 2018). Duration for both failed and successful campaigns both tend to be the same length of time and in this study, the duration has no effect on the state of a campaign. A study by (Mollick, 2014), the indicated that there is a relationship between duration and state of a campaign, but this fails to hold in this study as duration had no significant relationship between the state of campaign and success rate. The spotlight indicates how often the campaign creator has had other projects sponsored on the CFP as this would likely increase the success rate – since a communication frequency would have been created with backers.

Lastly, there is a positive impact between staff pick and success rate. This was explained in a study by (Lelo de Larrea et al., 2019) Employees might likely select projects they think are unique thereby attracting more backers. The state of projects campaign and success rate are both affected in two different ways with multiple factors.

7 Conclusion and Future Work

The success rate is what determines whether a project will likely exceed the funding target or reach the exact funding target. The result of this study indicated that some campaigns have a 0% success rate even when the funding target was low. It is imperative to say that factors that affect the state of a campaign are not also factors that determine the success rate. This is influenced by the platform and the campaign creator. In terms of the spotlight, there is a positive relationship between success rate and spotlight.

Crowdfunding as of today has had major impacts in the technological industry for funding massive campaigns. The study was centred around gaining a clear understanding of the factors affecting the performance of crowdfunding in agro-campaigns and offered recommendations to agrarian entrepreneurs. Crowdfunding is an alternative finance ameliorating the financial deficiency in agriculture. Mind you that this deficiency persists because of the low success rate of crowdfunding campaigns in agriculture. The agrarian entrepreneur needs to understand the mechanism of crowdfunding both offline and online to increase the success rate. A suggestion of further studies can be carried out using different crowdfunding platform that works with a different funding model to analyse how agrarian project creators perform.

.

References

.

Agrawal, A., Catalini, C. and Goldfarb, A., 2014. Some simple economics of crowdfunding. *Innovation policy and the economy*, *14*(1), pp.63-97.

Ahlers, G.K., Cumming, D., Günther, C. and Schweizer, D., 2015. Signaling in equity crowdfunding. *Entrepreneurship theory and practice*, *39*(4), pp.955-980.

Allison, T.H., Davis, B.C., Webb, J.W. and Short, J.C., 2017. Persuasion in crowdfunding: An elaboration likelihood model of crowdfunding performance. *Journal of Business Venturing*, *32*(6), pp.707-725.

Anshari, M., Almunawar, M.N., Masri, M. and Hamdan, M., 2019. Digital marketplace and FinTech to support agriculture sustainability. *Energy Procedia*, *156*, pp.234-238.

Awad, M., Khanna, R., 2015. Support Vector Machines for Classification. pp. 39–66.

Beaulieu, T., Sarker, S. and Sarker, S., 2015. A conceptual framework for understanding crowdfunding. *Communications of the Association for Information Systems*, *37*(1), p.1.

Belleflamme, P., Lambert, T. and Schwienbacher, A., 2013. Individual crowdfunding practices. *Venture Capital*, *15*(4), pp.313-333.

Bhadane, C., Dalal, H., Doshi, H., 2015. Sentiment Analysis: Measuring Opinions. Procedia Comput. Sci. 45, 808–814.

Carolan, M., 2017. Agro-digital governance and life itself: food politics at the intersection of code and affect. *Sociologia Ruralis*, *57*, pp.816-835.

Chang, W.I., 2016, December. Exploring Crowdfunding Performance of Agricultural Ventures: Evidence from FlyingV in Taiwan. In *International Conference on E-Infrastructure and E-Services for Developing Countries* (pp. 165-173). Springer, Cham.

Chemla, G. and Tinn, K., 2020. Learning through crowdfunding. *Management Science*, *66*(5), pp.1783-1801.

Chen, J., de Hoogh, K., Gulliver, J., Hoffmann, B., Hertel, O., Ketzel, M., Bauwelinck, M., van Donkelaar, A., Hvidtfeldt, U., Katsouyanni, K., Janssen, N., Martin, R., Samoli, E., Schwartz, P., Stafoggia, M., Bellander, T., Strak, M., Wolf, K., Vienneau, D., Vermeulen, R., Brunekreef, B. and Hoek, G., 2019. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Environment International*, 130, p.104934.

Cillo, V., Cardinali, S. and Bertoldi, B., 2017. CROWD-FOODING. The Role of Crowdfunding in the Food Industries Open Innovation In *10TH ANNUAL CONFERENCE OF THE EUROMED ACADEMY OF BUSINESS| Global and national business theories and practice: bridging the past with the future* (pp. 387-404). EuroMed Press.

Cumming, D.J., Leboeuf, G. and Schwienbacher, A., 2015. Crowdfunding models: Keep-it-all vs. all-or-nothing. *Financial Management*.

Dhakal, C.P., 2017. Dealing with Outliers and Influential Points while Fitting Regression. J. Inst. Sci. Technol. 22, 61–65. https://doi.org/10.3126/jist.v22i1.17741

Doreswamy,Vastrad, C.M., 2013. Performance Analysis of Neural Network Models for Oxazolines and Oxazoles Derivatives Descriptor Dataset. *International Journal of Information Sciences and Techniques*, 3(6), pp.1-15.

Ellman, M. and Hurkens, S., 2019. Optimal crowdfunding design. *Journal of Economic Theory*, *184*, p.104939.

Etter, V., Grossglauser, M., Thiran, P., 2013. Launch hard or go home!: predicting the success of kickstarter campaigns, in: Proceedings of the First ACM Conference on Online Social Networks - COSN ’13. Presented at the first ACM conference, ACM Press, Boston, Massachusetts, USA, pp. 177–182. https://doi.org/10.1145/2512938.2512957

Farhadloo, M., Rolland, E., 2016. Fundamentals of Sentiment Analysis and Its Applications. pp. 1–24. https://doi.org/10.1007/978-3-319-30319-2_1

Filimonova, N.G., Ozerova, M.G., Ermakova, I.N. and Miheeva, N.B., 2019, August. Crowdfunding as the way of projects financing in agribusiness. In *IOP Conference Series: Earth and Environmental Science* (Vol. 315, No. 2, p. 022098). IOP Publishing.

Gerber, E.M., Hui, J.S. and Kuo, P.Y., 2012, February. Crowdfunding: Why people are motivated to post and fund projects on crowdfunding platforms. In *Proceedings of the international workshop on design, influence, and social technologies: techniques, impacts and ethics* (Vol. 2, No. 11, p. 10). Northwestern University Evanston, IL.

Hemer, J., 2011. *A snapshot on crowdfunding* (No. R2/2011). Arbeitspapiere Unternehmen und Region.

Hildebrand, T., Puri, M. and Rocholl, J., 2017. Adverse incentives in crowdfunding. *Management Science*, *63*(3), pp.587-608.

Hornuf, L. and Schwienbacher, A., 2017. Should securities regulation promote equity crowdfunding?. *Small Business Economics*, *49*(3), pp.579-593.

Hsieh, F.Y., Lavori, P.W., Cohen, H.J., Feussner, J.R., 2003. An Overview of Variance Inflation Factors for Sample-Size Calculation. Eval. Health Prof. 26, 239–257. https://doi.org/10.1177/0163278703255230

Hui, J.S., Greenberg, M.D., Gerber, E.M., 2014. Understanding the role of community in crowdfunding work, in: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW ’14. Association for Computing Machinery, Baltimore, Maryland, USA, pp. 62–74. https://doi.org/10.1145/2531602.2531715

Huang, C.C., 2020. Cocreating social innovations between an agro-food company and rice farmers in Taiwan: exploring the process mechanisms. *British Food Journal*.

Khalatur, S., 2017. Finding of alternative sources of innovations funding in agriculture. *Technology audit and production reserves*, (3 (5)), pp.20-25.

Korkmaz, M., GÜNEY, S. and YİĞİTER, Ş.Y., 2012. The importance of logistic regression implementations in the Turkish livestock sector and logistic regression implementations/fields.

Kuppuswamy, V. and Bayus, B.L., 2018. Crowdfunding creative ideas: The dynamics of project backers. In *The economics of crowdfunding* (pp. 151-182). Palgrave Macmillan, Cham.

Lelo de Larrea, G., Altin, M., Singh, D., 2019. Determinants of success of restaurant crowdfunding. Int. J. Hosp. Manag. 78, 150–158. https://doi.org/10.1016/j.ijhm.2018.10.003

Levy, R., Navereau, B. and Triboulet, P., 2018. The Trajectory of Innovative Collaborative Projects in Agro-food Sector: the Role of Proximities and Intermediaries in Knowledge Flow. *Policy*, *35*, p.23.

Lukkarinen, A., Teich, J.E., Wallenius, H. and Wallenius, J., 2016. Success drivers of online equity crowdfunding campaigns. *Decision Support Systems*, *87*, pp.26-38.

Luo, T., Chen, S., Xu, G., Zhou, J., 2013. Sentiment Analysis. pp. 53–68. https://doi.org/10.1007/978-1-4614-7202-5_4

Mariani, A., Annunziata, A., Nacchia, F. and Vastola, A., 2016. Crowdsourcing in wine business: Co-creation and fundraising experiences. In *Successful social media and ecommerce strategies in the wine industry* (pp. 93-114). Palgrave Macmillan, New York.

Markas, R. and Wang, Y., 2019. Dare to Venture: Data Science Perspective on Crowdfunding. *SMU Data Science Review*, *2*(1), p.19.

Marill, K.A., 2004. Advanced statistics: linear regression, part II: multiple linear regression. *Academic emergency medicine*, *11*(1), pp.94-102.

Mitra, D., 2012. The role of crowdfunding in entrepreneurial finance. *Delhi Business Review*, *13*(2), pp.67-72.

Mollick, E., 2014. The dynamics of crowdfunding: An exploratory study. *Journal of business venturing*, *29*(1), pp.1-16.

Moritz, A. and Block, J.H., 2016. Crowdfunding: A literature review and research directions. In *Crowdfunding in Europe* (pp. 25-53). Springer, Cham.

Ogutu, J.O., Schulz-Streeck, T., Piepho, H.-P., 2012. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. BMC Proc. 6, S10. https://doi.org/10.1186/1753-6561-6-S2-S10

Pacchi, C. and Pais, I., 2020. Civic Crowdfunding in Milan: Between Grass-Roots Actors and Policy Opportunities.

Pronti, A. and Pagliarino, E., 2019. Not Just for Money. Crowdfunding a New Tool of Open Innovation to Support the Agro-Food Sector. Evidences on the Italian Market. *Journal of Agricultural & Food Industrial Organization*, *17*(1), pp.1-16.

Schwienbacher, A. and Larralde, B., 2010. Crowdfunding of small entrepreneurial ventures. *Handbook of entrepreneurial finance, Oxford University Press, Forthcoming*.

Swai, D., 2019. *Crowd funding in Tanzania as a source of fund to businesses looking at opportunities, challenges and solutions* (Doctoral dissertation, Mzumbe University).

Thaker, H.M.T., Khaliq, A., Sakaran, K.C. and Thaker, M.A.M.T., 2020. A discourse on the potential of crowdfunding and Islamic finance in the agricultural sector of East Java, Indonesia. *Jurnal Ekonomi & Keuangan Islam*, *6*(1), pp.10-23.

Uyanık, G.K. and Güler, N., 2013. A study on multiple linear regression analysis. *Procedia-Social and Behavioral Sciences*, *106*, pp.234-240.

Vulkan, N., Åstebro, T. and Sierra, M.F., 2016. Equity crowdfunding: A new phenomena. *Journal of Business Venturing Insights*, *5*, pp.37-49.

Wang, W., Zhu, K., Wang, H. and Wu, Y.C.J., 2017. Impact of sentimental factor on the successful Crowdfunding campaigns: a text mining approach.

Xu, B., Zheng, H., Xu, Y. and Wang, T., 2016. Configurational paths to sponsor satisfaction in crowdfunding. *Journal of Business Research*, *69*(2), pp.915-927.