

Predictive analysis for calculating the valuation of the affiliated fund of a private pension system using machine learning techniques and tools

Jimmy Armas Aguirre¹, Jhonatan Espinoza Ladera¹, Brian Dueñas Castillo¹, Santiago Aguirre Mayorga²

¹ Universidad Peruana de Ciencias Aplicadas, Lima, Perú

jimmy.armas@upc.pe, u201214060@upc.edu.pe, u201214711@upc.edu.pe

² Pontificia Universidad Javeriana, Bogotá, Colombia

saguirre@javeriana.edu.co

Abstract - this paper proposes a model for the analysis of the prediction of the accumulated fund for affiliates based on an area of study such as machine learning. The model allows to predict the pension fund of an affiliate in the private pension system by means of a web solution. In this sense, people will have, information and an adequate tool that allow them to have an oversight of the valuation of their funds throughout the years until retirement. In Peru, the decree of law 1990 states that the age for retirement is 65 years old, although there is also the option for early retirement. The proposed model consists of data analytics usage based on the modeling of machine learning algorithms through cloud platforms. The model structure includes four layers: transformation of the affiliate's data, security and privacy of personal data, obtaining and managing data, and finally, the life cycle of data applied to analytics. The model emphasizes data analytics concepts where large amounts of data are examined that lead to conclusions for better decision making. In doing this, the machine learning technique "boosted decision tree" is used due to the proximity of this technique applied in the financial forecast. The model was validated with a Pension Fund Administrator (AFP) in Lima (Peru) and the results obtained focused on the use of improved decision tree regression with a coefficient of determination of 99.997% and an average square error of 0.00650%. The coefficient of determination is an indicator of the quality of the model to predict results while the quadratic error quantifies the percentage of error among the set of results obtained under the boosted decision tree regression model.

Keywords— predictive models, AFP Fund, pension fund administrator, predictive analysis, machine learning, decision trees.

I. INTRODUCTION

The lack of knowledge about how personal finances are managed in private pension funds represents a risk in the management of each individual's wealth. Currently, there are several methods for calculations of the pension fund where pension fund managers simulate the accumulated fund and the pension that a retiree will receive. However, what is not offered is the possibility of forecasting the fund from growth factors based on the affiliate's own criteria that give them the opportunity to choose which voluntary pension fund can be the most beneficial at the end of their working life, knowing at all times what they are exposed to.

By 2017, only 16% of Peruvians (representing approximately

4.6 million of the population) were informed of the amount they will receive as a pension when they retire [1]. Also, the official information that is accessed is limited, unreliable and unintelligible to a regular person. This limited knowledge of personal finances is also due to poor financial education in Peru: less than 5% consider having financial knowledge [2].

On the other hand, the current system is based on methodologies of complex calculation for the forecast of the accumulated fund of each affiliate and the factors that govern such calculations use standard growth factors for the entire population [3]. For what to consider a personal aspiration of salary and personal growth, certain people work well, for the Pareto of the population that earns the most in the country is inapplicable under the current methodology.

The technology has evolved and great advances can be seen with respect to the capture and storage of information. One of these technologies uses cloud computing, which has grown in large proportions in various industries and attracts the attention of research communities due to its potential to implement customized technological solutions [4].

This paper proposes a model of predictive analysis supported by a cloud technology platform based on techniques and machine learning tools to determine the accumulated fund of an affiliate that contributes to a private pension system.

This model combines a series of variables belonging to an affiliate, such as age, current salary, level of education, gender, place of residence, years of contribution and occupation. The objective of this model is to allow the affiliate to use their own growth variables that allow them to forecast their fund under a customized analysis for better decisions in the management of their personal finances and, complementing the information provided by the AFPs, to achieve their personal and family finances goals.

This paper presents the following structure: starts with a literature review, and the second part presents the predictive model. Then, the preliminary results of the case study are described and analyzed and finally the conclusions and future work is presented.

II. LITERATURE REVIEW

There are previous studies that investigate issues presented into the private pension system using mathematical models, however, there are only few documents that perform projection exercises based on statistical models and using technology as support: Belolipetskii and Lepskaya propose a stochastic model of the variation of the composition by age of the members of the fund [11], which describes a pension fund with any entry / exit condition applying methods of actuarial mathematics, predicting the future financial flows of the fund and its

Digital Object Identifier (DOI):

<http://dx.doi.org/10.18687/LACCEI2019.1.1.343>

ISBN: 978-0-9993443-6-1 ISSN: 2414-6390

dynamics of accumulation. Alonso and Tuesta, document a model of macroeconomic and actuarial projection to simulate the expected coverage for the period 2012 - 2050 [8], the projection model uses databases with information on the socioeconomic profiles of almost 4 million people, with 40 variables in the period 2010-2050.

The data was organized into 60 types of representative individuals, which are then projected following a generational accounting model based on the use of actuarial techniques and macroeconomic estimates for projection scenarios. These permanent characteristics will determine key elements such as the membership rate, the contribution rate, the contribution salary, etc. The model also includes a dynamic element: the education variable that affects the probability of contribution, salary profiles and contribution densities.

In Peru, contributions in pension funds are compulsory by law. For the AFP system the amount corresponds to 10% of the taxable income, where it will be invested into three different funds depending on the risk that individual wants to assume. These range from fund 0 to fund 3, where the first corresponds to the riskiest and the last to the one that assumes the least risk. For this work, we define predictive capacity as the ability to construct and evaluate models to generate predictions of new observations, using sophisticated statistical tools to identify what might happen in the future [5].

There are machine learning techniques that are applied to solve problems of different nature, but none have been identified as the best over the others since it is a very complex task and many of the existing studies have not been carried out with the care that is required. [6]

To create predictive capacity, organizations have to rely on a predictive analytics platform that incorporates data warehouses, predictive analytics algorithms (regression analysis, machine learning, and neural networks), dashboards, and presentation of reports that provide information for optimal decision making. For the evaluation and application of machine learning techniques it is necessary to have a considerable amount of data. The greater the amount of data, the accuracy of the predictions increases [7].

A. Cloud platforms evaluation

We decided to use Azure ML due to the scalability, traceability and performance and model varieties that incorporates. It also offers several distinctive features, such as version control, modeling, easy operation and user code integration by simplifying the process of predictive models. Below is Table I with a brief description of the cloud platforms that are in the market:

Table I. Cloud platforms for machine learning

Features	Amazon	Google ML	Azure
Cloud Service Area	IaaS,MLaaS [26]	PaaS,MLaaS [26]	IaaS,Paas,MLaaS [26]
Development Tools	Only provides virtual machines to create images of server	Provides editing, simulation and deployment tools. [25]	Provides “Visual Studio” which is one of the IDE for developing cloud-based projects using

	platform. [24]		Microsoft Technology. [23]
Programming Language Support	Supports Oracle and .NET.	Python and Django.	NET, JAVA, Apache, Ruby, PHP, C and C++.
Algorithms	Classification, regression and clustering.	Classification, regression, clustering, anomaly detection, recommends and rankings.	Classification, regression and recommends.
Benefits	Extensive and mature offering. Enterprise-friendly services. Open and flexible global reach.	Open source support and portability. Discounts & flexible contracts. Designed for cloud-based businesses.	Integration with Microsoft tools. Broad feature set. Ranks first in development and testing tools. Open source support. Hybrid cloud-
Pricing	Per hour	Per minute	Per minute
Availability	99.95% monthly availability. [27]	99.95% monthly availability. [28]	99.95% monthly availability. [29]

B. Algorithm evaluation

Table II describes the types of algorithms that can be used in this type of studies to the proposed model uses regression algorithms since it focuses on the predictions of continuous numbers, also based on forecasting from one or more causes (independent variables), which may be for example time, product or service prices, prices of competition, economy of the country, government actions or advertising promotions.

Table II. Types of algorithms

Types of algorithms	Description	Algorithms
Classification algorithms	The classification algorithms are used when the desired output is a discrete label. In other words, they are useful when the answer to your question about your business is within a finite set of possible outcomes. Many	Boosted decision tree, Decision jungle, Locally deep SVM, Logistic

	use cases, such as determining whether an email is spam or not, only have two possible outcomes. This is called binary classification [19]	regression, Neural network [19].
Regression algorithm	Regression algorithms are used to predict variables with numerical results, some techniques used are linear regression, decision trees and neural networks that mimic the human brain [20]	Bayesian linear regression, Boosted decision tree regression, Linear regression, Decision forest regression, Random forest regression, Neural network, Ordinal regression, Poisson regression [20].
Clustering Algorithms	Grouping algorithms are designed to explore an inherent natural structure of data objects, where objects in the same group are as similar as possible and objects in different groups are as different as possible. The equivalence classes induced by the groups provide [21]	K-means clustering [21].

C. Variables for the prediction of the accumulated fund

Several factors have been identified that make it possible to show which variables influence the projection of an accumulated fund, which are shown in this section. Table III shows 15 of the most significant variables found in the literature review. From these variables, it is necessary to establish the relevant ones that are required to project the value of the accumulated fund of a person in a private pension system.

Table III. Main variables for the projection of the accumulated fund.

Factor	Description	Reference
Age	Affiliate age	[11], [8], [9], [10], [12], [13]
Gender	Affiliate gender	[9], [10], [13], [8].

Salary	Current compensation received by the affiliate.	[14], [10], [8],
Civil status	Marital status of the member (single, married, divorced).	[8], [14]
Education degree	Affiliate study level.	[8], [14]
Macroeconomic variable	Factors that influence the profitability of the fund.	[11], [12], [8], [13]
Salary growth rate	Percentage rate of salary increase.	[9], [10], [11], [13]
Discount rates	Discount rate of the pension fund.	[10], [9]
Life expectancy	Life expectancy of the affiliate.	[10], [14],
Place of residence	City where the affiliate resides.	[10]
Number of years of contribution	Number of years of contribution of the affiliate.	[11], [12]
Job Position	Current job position (Manager, Specialist, Analyst).	[14]

D. Algorithms for the prediction of the accumulated fund

As shown on table 5, the regression algorithms achieved better results like the correlation between the variables and expected error. A regression modeling study was conducted to identify effects on retirement. These algorithms were tested in an experiment using the Azure Machine Learning Studio platform to analyze, quantify and compare the results.

Table IV. Metrics and evaluation results of regression algorithms

Algorithm	Absolute Error Medium	Mean Square Error	Determination Coefficient R ²	Absolute Relative Error
Linear regression	0.00543	0.006535	0.805516	0.457369
Regression in neural networks	0.005628	0.006752	0.792346	0.474076
Boosted decision tree regression	0.000065	0.000088	0.999965	0.005453
Regresión lineal bayesian	0.00543	0.006535	0.805517	0.457370
Decision forest regression	0.000022	0.000109	0.999946	0.001886

E. Prediction techniques

The regression algorithm boosted decision tree, obtained a better performance with a higher correlation between the objectives and expected results. A regression model study was conducted to identify retirement effects [15]. Data from the labor force survey was used to implement a discontinuity-acute regression framework, in which the variable in execution is the distance from the month of individual birth to a certain reference month (See Table IV).

III. PROPOSED MODEL: PREDICTIVE ANALYSIS

A. Model description

The model proposes a predictive analysis for a pension fund through the use of data analytics and services in the cloud. This model serves for the affiliate to project their pension fund in a reliable and personalized manner according to their own economic levels and aspirations for growth. This results in better financial knowledge of an affiliate's funds for the future. The design contemplates four layers and is based on the use of technologies in the cloud.

The first layer implicates the process of transforming the affiliate's data. The second access the security and privacy of affiliate's personal data. The third, the collection and management of data as the set of activities that convert them into data with real value. Finally, the fourth layer: allows the management of the data life cycle. Likewise, the model contemplates a technological architecture proposal where the technological elements that support the model are included.

In figure 1, the phases, components and elements that abstractly represent the proposed model are shown. Likewise, the layers of security, privacy, procurement, management and life cycle of the data to be used in the model are represented.

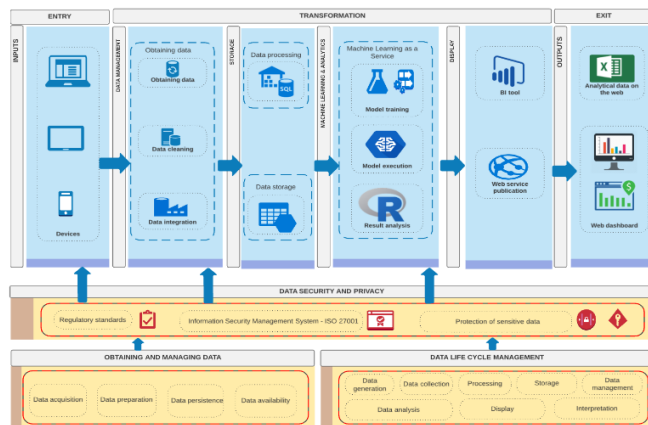


Figure 1. Predictive analysis model of the accumulated affiliated fund in the private pension system

- **Affiliate's data transformation layer**

The use of web applications in smartphones, notebooks, tablets, desktops, etc., and the incessant increase of users accessing the internet, generate a considerable volume of online information that will be stored, processed and analyzed under the proposed model. Each phase of the model has a specific objective in the data processing and depends on the results of each previous phase. This leads to presenting a dynamic model, which ensures

the good decision making of the affiliates in the management of the future of their funds.

This layer has the following components:

Input data. Information entered by the user through any mobile device, computers, tablets, etc.

Data management. Obtaining sources of information (relational databases); data validation and cleanup process; and grouping and integration of information.

Storage. Loading and processing of information for storage in a data collection invariable in time.

Machine learning & analytics. Application and use of machine learning cloud services for the training, execution and evaluation of the prediction model.

Output data. Consumption of web services through the use of mobile web applications.

- **Data security and privacy layer.**

When using personal data of the affiliate as input data, it should be taken into consideration the following regulatory aspects:

ISO 27001 - Information security management system.
Protection of sensitive data.

- **Obtaining layer and data management.**

As part of a data maintenance process that supports the analytical process of the first phase, it is necessary to carry out the following activities:

Data acquisition.
Data preparation
Data persistence
Data availability

- **Data life cycle management layer.**

In order to ensure a correct predictive analysis and obtain results not very far from reality, the data life cycle consists of:

Generation of data
Harvest
Processing
Storage
Data management
Analysis
Display
Interpretation

B. Technological architecture model

The architecture to be used supports the predictive analysis model of the AFP fund during the affiliate's consultation by means of an application. The architecture shows the technological support for web applications based on the C # language and using visual studio as an integrated development environment (IDE).

For its implementation, Microsoft IIS (Internet Information Services) is used as a web server and SQL Server to manage the database. This architecture is adaptable to changes according to the need of the application as shown in Fig. 2.

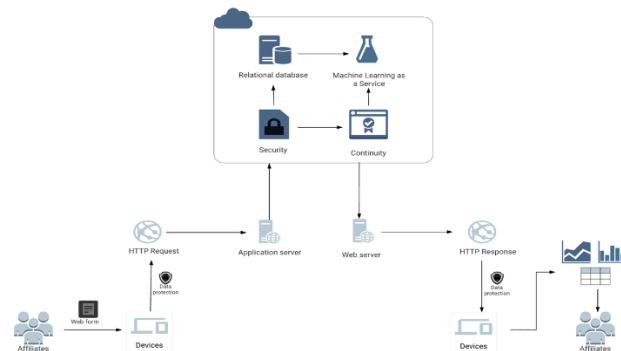


Figure 2. Technological architecture for the predictive analysis of the accumulated affiliated fund in the Private Pension System

The use of the Visual Studio development environment with which the application and the interface is created using CSS, HTML5 and Javascript is shown. The registration or entry of the user will have validation measures related to data security. To generate the model, obtain results and generate reports, the tools of Azure Machine Learning and Power BI are used. The technology of Microsoft Azure to be in the cloud reduces the complexity, processing and infrastructure in the construction of the model. But this does not exempt him from having knowledge in application development, managing databases, managing BI tools (Business Intelligence) and knowledge of cloud services. All this in order to create a solution that supports the model giving as results a prediction with a high credibility range.

IV. CASE STUDY

A. Organization

An AFP pension fund management entity in Peru was the company that provided the database with information from 2010 to 2015 in order to validate the proposed model. Currently, they have a network of 15 agencies nationwide and a portfolio of approximately more than 1.6 million affiliates. Prior to the validation of the proposed model, there was no personalized predictive analysis tool to project the accumulated fund of each affiliate. Therefore, a web application supported by the predictive analysis model was developed in order to test and obtain results that help validate the project. For this, the entity facilitated the delivery of test data that allowed to simulate and validate the model with results very close to reality. We worked with information on 45,600 affiliates of legal age residing in different parts of Peru who contributed to the system during the years 2010-2015.

B. Implementation

• Input data

Affiliate data necessary for authentication at the time of using the web application: each affiliate is registered, the terms and conditions on the protection of personal data are accepted and web security methods are used. Also, it is recommended that the device has a web browser. In Fig. 3 the user authentication view and the registration view for new user are shown.

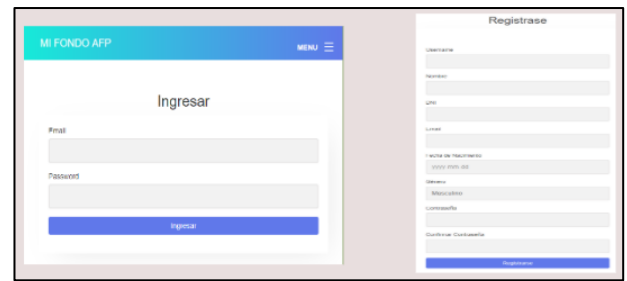


Figure 3. Registration screen of the My Fund AFP application

• Data management

The web application records the input data through a web form that is immediately sent to the azure machine learning cloud service. The web application registers the input data through a web form and are sent immediately to the Azure Machine Learning cloud service.

• Storage

The cloud storage service reduces the complexity of inputting and storing the information of affiliates before carrying out all types of processing and data analysis in batches. It adapts to the Azure platform and the different object-oriented programming languages. Additionally, the same data is stored in a SQL Database Manager in order to have a backup of the information.

• Data pre-processing

After the storage process of the database set, the first thing is to have it clean, without missing values, nulls or atypical data. For this reason, the pre-processing of the data is fundamental to be able to predict the accumulated fund of an affiliate according to its characteristics. Figure 4 shows the pre-processing steps.

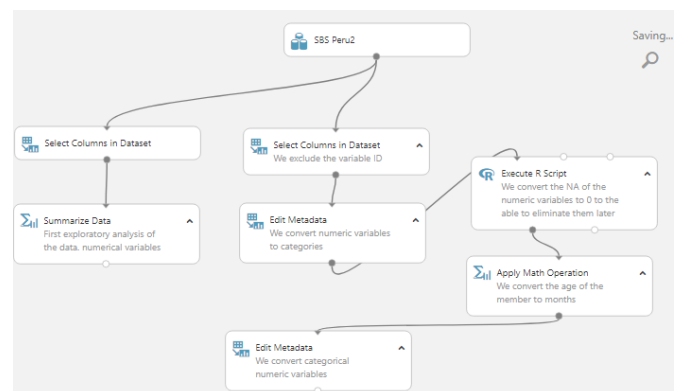


Figure 4. Execution of data pre-processing

• Variable selection

After data pre-processing, the data set was analyzed to find the variables that has a greater predictive power or degree of correlation for the response variable. We have used three azure ml methods to detect the variables with more predictive power (Fig. 5). For this, the "filter based feature selection" function was chosen.

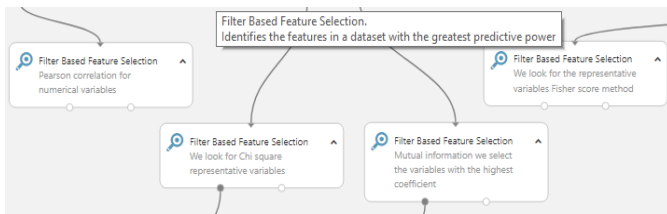


Figure 5. Regression algorithms execution

- Machine learning and analytics

Five machine learning models were evaluated with the selected database, to determine which of the models best fits the actual accumulated fund. The models implemented are: linear regression, regression in neural networks, boosted decision tree regression, bayesian linear regression and decision forest regression. All models were evaluated by cross-validation with the exception of neural networks that made a comparison between cross-validation and normal validation. Figure 6 shows the training and evaluation steps of the algorithms.

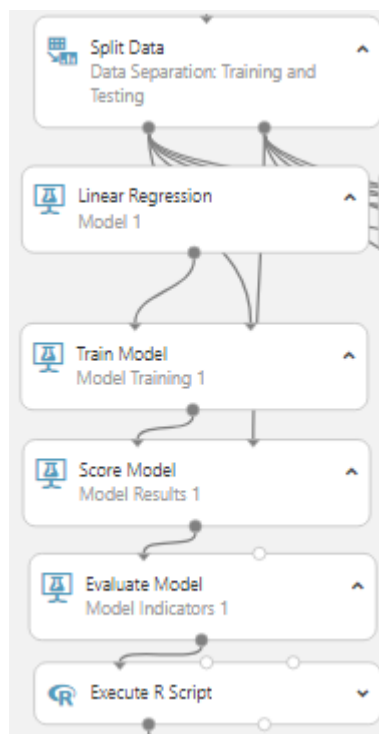


Figure 6. Training and evaluation of regression algorithms.

- Evaluation of the algorithms

The accuracy and effectiveness rate were compared for each algorithm under five efficiency indicators: mean absolute error (MAE), mean square error (RMSE), relative absolute error (RAE), relative square error (RSE) and the coefficient of determination (R2) (See Figure 7). This preliminary information gives the private entity the ability to understand the actual behavior of the accumulated fund based on each unique combination of characteristics of the set of affiliates, which leads to obtaining different projections of the fund for each profile of an affiliate.

Algorithm	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Relative Squared Error	Coefficient of Determination
Linear Regression	0.00543	0.006535	0.457369	0.194484	0.805516
Neural Network Regression	0.005628	0.006752	0.474076	0.207654	0.792346
Boosted Decision Tree Regression	0.000065	0.000088	0.005453	0.000035	0.999965
Bayesian Linear Regression	0.00543	0.006535	0.45737	0.194483	0.805517
Decision Forest Regression	0.000022	0.000109	0.001886	0.000054	0.999946

Figure 7. Results of the comparative analysis of the regression algorithms.

- Output data

In Fig. 8, a view of the web application is shown where the results of the simulation are presented by calculating the current methodology versus the projection of the accumulated fund using Azure ML (See Table I). After the personal data have been entered, the tool allows calculating the projection of the fund, the final accumulated fund at the time of retirement of the affiliate and the pension that the person will receive. These results are shown in two graphs that overlap in the same two-dimensional plane. On the one hand, the projection is plotted according to the current methodology and on the other hand the projection using the proposed machine learning model.

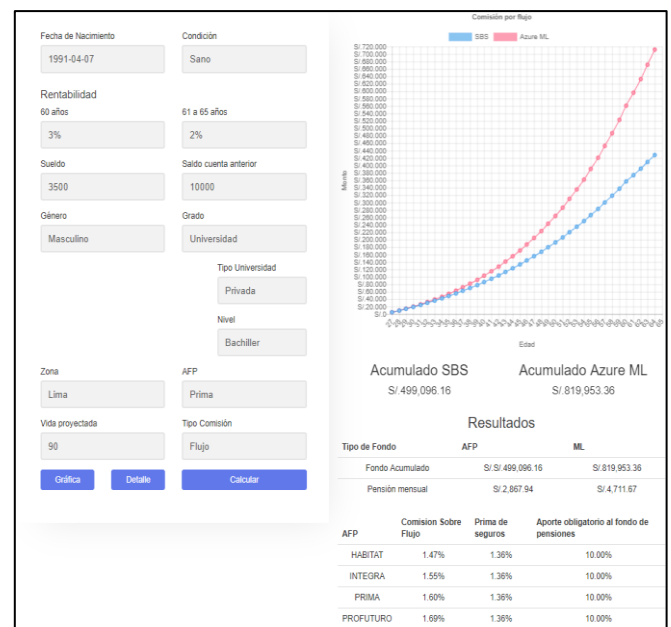


Figure 8. Projection of the accumulated fund using the current methodology vs machine learning in the cloud

C. Results

According to the data obtained in Fig. 7 and Fig. 8, the importance of predictive analytics in the projections of pension funds in favor of more real information to the affiliate for better management of their finances is checked. The analysis of data in the cloud facilitates the calculation and projection of the accumulated fund in real time according to the user's query.

Within the algorithms chosen to model and validate the credibility of the results, it was observed in Table III that modeling based on Boosted Decision Tree Regression gives better results compared to the other four models, since it offers the highest R^2 (99,997. %), a minimum RMSE with respect to the others (0.00000880%) and a lower RSE (0.00650%). This is the regression algorithm used to implement a web service that allows estimating and projection of the accumulated fund of an affiliate to a private pension system. These indicators reflect the benefit of the adjustment of the model to the variable that we intend to predict, taking into account that the result of the determination coefficient oscillates between 0 and 1. While closer to 1 the value is placed, the greater the adjustment of the model to the variable that we are trying to explain. Conversely, the closer to zero, the less fit the model will be and, therefore, the less reliable it will be.

The size and quality of the data used were very close to reality, so the results delivered to the private entity represent financial projections of high credibility.

V. CONCLUSIONS

In this work, we have proposed a model for the predictive analysis of the accumulated affiliated fund to the private pension system. The model could be used on any mobile device with internet connection without the need to mount an entire infrastructure or technological platform since all resources are consumed and processed from the cloud platform.

The model was validated in an AFP entity in Lima, Peru. The results based on the bayesian regression, linear reversal, decision forest regression, boosted decision tree regression and neural networks modeling showed that the boosted decision tree regression technique obtained a coefficient of determination of 99.997% with an average square error of 0.00650%. These percentages were calculated based on the test data provided by the private entity AFP demonstrating a minimum error percentage in the test results to project the accumulated fund of an affiliate. In the same way, the obtained results quantitatively validate the level of proximity for each applied technique and the differences of the variation of the accumulated fund of the affiliate comparing it with the current methodology of calculation. Both the AFP entity and the affiliates showed satisfaction in the design of the web tool due to its simplicity, good interaction and friendly graphics.

Finally, the affiliate can simulate the fund's projection under different scenarios and according to its own growth factors. This will allow the affiliate to plan, control and make better decisions about the future of their finances in order to achieve personal and family well-being. On the other hand, the AFP entities verified the high impact of the model as a predictive analysis tool complementary to other financial models already applied in their processes.

REFERENCES

- [1] Ahorro LATAM 360°, «SBS» [Online]. Disponible: http://www.comoahorramos.com/Estudio_Ahorro_Latam_360_reducida.pdf
- [2] Konicz, A.K., & Mulvey, J.M. (2015). Optimal savings management for individuals with defined contribution

- pension plans. *European Journal of Operational Research*, 243, 233-247.
- [3] Resolución SBS N° 9617-2012 (Anexo Técnico 1), «SBS» [Online]. Disponible: <https://www.sbs.gob.pe/app/Calculadora/Paginas/metodologia.pdf>
- [4] M. Ali, S.U. Khan, A.V. Vasilakos, «Security in Cloud Computing: Opportunities and Challenges, *Information Sciences*, 2015», doi: <http://dx.doi.org/10.1016/j.ins.2015.01.025>.
- [5] Yao, Y., Xiao, Z., Wang, B., Viswanath, B., Zheng, H., & Zhao, B.Y. (2017). Complexity vs. performance: empirical analysis of machine learning as a service. *IMC*.
- [6] Ribeiro, M.C., Grolinger, K., & Capretz, M.A. (2015). MLaaS: Machine Learning as a Service. 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), 896-902.
- [7] Lee et al. (2019). Global stock market investment strategies based on financial network indicators using machine learning techniques. 228-242. 10.1016/j.eswa.2018.09.005
- [8] Alonso, Javier & Hoyo, Carmen & Tuesta, David. (2015). A model for the pension system in Mexico: diagnosis and recommendations. *Journal of Pension Economics and Finance*. 14. 76-112.
- [9] A. Belolipetskii, A & A. Lepskaya, M. (2018). A Mathematical Model of Pension Fund Operation and Methods of Fund Stability Analysis. *Computational Mathematics and Modeling*. 29. 10.1007/s10598-018-9404-7.
- [10] M. de Klerk, Antonie. (2017). Decision Making under Uncertainty and Risk: A Retiree's Perspective on Choosing between Pension Options. 1-12. 10.23919/PICMET.2017.8125309.
- [11] Kim, Hyungsu & Kim, Geonwoo & Lee, Sungchul. (2016). THE ANALYTIC APPROACH FOR THE STOCHASTIC PROJECTION OF THE PUBLIC PENSION FUND. *Probability in the Engineering and Informational Sciences*. 31. 1-11. 10.1017/S0269964816000504.
- [12] Stancanelli, Elena. (2017). Couples' Retirement under Individual Pension Design: a Regression Discontinuity Study for France. *Labour Economics*. 49. 10.1016/j.labeco.2017.08.009.
- [13] Badreddine, El Goumi & El Khomssi, Mohammed & Fikri, Majda. (2016). Model for the management of pension fund with deterministic and stochastic parameters. 1-5. 10.1109/GOL.2016.7731665.
- [14] Vonkova, Hana & van soest, Arthur. (2014). How Sensitive Are Retirement Decisions to Financial Incentives: A Stated Preference Analysis. *Institute for the Study of Labor (IZA), IZA Discussion Papers*. 29. 10.1002/jae.2313.
- [15] Ripley, Brian. (2018). MASS: Support Functions and Datasets for Venables and Ripley's MASS. <https://cran.r-project.org/web/packages/MASS/index.html>
- [16] Osnabrugge et al. (2018). Prediction of Costs and Length of Stay in Coronary Artery Bypass Grafting. 1286-1293. 10.1016/j.athoracsur.2014.05.073
- [17] Breiman. (2001). RandomForest: Breiman and Cutler's Random Forests for Classification and Regression. 5-32. 10.1023/A:1010933404324
- [18] Maechler et al. (2018). Robustbase: Basic Robust Statistics. 10.2139/ssrn.3177507
- [19] Hai, Mo & Zhang, You & Zhang, Yuejin. (2017). A Performance Evaluation of Classification Algorithms for

- Big Data. *Procedia Computer Science*. 122. 1100-1107. 10.1016/j.procs.2017.11.479.
- [20] Kibekbaev, Azamat & Duman, Ekrem. (2015). Benchmarking Regression Algorithms for Income Prediction Modeling. 180-185. 10.1109/CSCI.2015.162.
- [21] Nerurkar, Pranav & Shirke, Archana & Chandane, Madhav & Bhirud, Sunil. (2018). Empirical Analysis of Data Clustering Algorithms. *Procedia Computer Science*. 125. 770-779. 10.1016/j.procs.2017.12.099.
- [22] Yao, Y., Xiao, Z., Wang, B., Viswanath, B., Zheng, H., & Zhao, B.Y. (2017). Complexity vs. performance: empirical analysis of machine learning as a service. *IMC*.
- [23] MICROSOFT AZURE (2019) – What is Azure Machine Learning Studio? [online] Disponible: <https://docs.microsoft.com/en-us/azure/machine-learning/studio/what-is-ml-studio>
- [24] AMAZON ML (2019) – ¿Qué es Amazon Machine Learning? [online] Disponible: <https://docs.aws.amazon.com/es-es/machine-learning/latest/dg/what-is-amazon-machine-learning.html>
- [25] GOOGLE CLOUD (2019) – Análisis Predictivo: Cloud Machine Learning Engine [online] Disponible: <https://cloud.google.com/ml-engine/?hl=es-419>
- [26] Ribeiro, M.C., Grolinger, K., & Capretz, M.A. (2015). MLaaS: Machine Learning as a Service. 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), 896-902.
- [27] Amazon Web Services (2018)- Contrato de Nivel de Servicios de Computación de Amazon [online] Disponible: https://d1.awsstatic.com/legal/amazon-ec2-sla/Amazon_EC2_Service_Level_Agreement_-_Spanish_Translation_2018-02-12.pdf
- [28] Google CLOUD SLA (2018)- Google Compute Engine Service Level Agreement (SLA) [online] Disponible: <https://cloud.google.com/compute/sla>
- [29] Azure Microsoft SLA (2018) - Resumen de acuerdos de nivel de servicio [online] Disponible: <https://azure.microsoft.com/es-es/support/legal/sla/summary/>