# Predicting Chilean Retirement Plan Decisions with Deep Neural Networks

**Sebastián Espinoza**
Stanford GSB

**Akshay Malhotra**
Stanford GSB

**Daniel Zhang**
Stanford Engineering

## Abstract

Retirees in Chile spend a considerable portion of their savings on financial advisors to select the correct financial product for their retirement income. We apply a deep neural network approach to see how well we can approximate these decisions using publicly available retirement data, merged with other socioeconomic and demographic data for Chilean communities. With limited access to personal information, particularly data on risk aversion or spending patterns, we are able to achieve a reasonably good level of predictive accuracy, much higher than generally achieved in social science research, thus offering significant promise for future research and innovation in this space.

## 1 Introduction

In Chile, people face a very important financial decision at the time of their retirement in terms of how they want to receive and use their savings over the course of the rest of their lives. Once men turn 65 or women turn 60, they are presented with the following four retirement options:

1. Immediate annuity program
2. Temporal programmed withdrawal with deferred annuity program
3. Immediate annuity program with programmed withdrawal
4. Programmed withdrawal

This is a fairly substantial decision that depends on many factors surrounding the retiree's health, style of living, behavior, and other aspects of his or her life. In fact, according to the *Wall Street Journal*, the average pensioner receives slightly over a third of his or her pre-retirement income from a pension [Dube (2016)].

In order to properly weigh the pros and cons of each option, Chileans often consult retirement advisors, who can charge quite a bit for their services. Since 2008, such consultations have cost up to 2% of retirees' total savings, and this is by no means a small sum for them to pay. We're interested in understanding retirees' behaviour when deciding what to do with their savings. Hence, the project's goal is to, for the first time, predict with deep neural networks the retirement plan choice and evaluate if the model's accuracy resembles the option of an expensive advisor.

What makes this problem particularly interesting is the limited available data. In particular, we aim to achieve a reasonable level of accuracy with limited data on an individual's risk aversion or personal tastes; most our data is on a more macro-level, and for this reason, even a moderate level of success

,

will open many new doors for further research and innovation in this space. Generally, predictive models in social sciences have much lower accuracy levels than other more common machine learning models because it's extremely difficult to predict human behavior, but we nevertheless try to explore this problem due to its immense social value.

## 2 Related work

Although there isn't any literature on machine learning (ML) techniques applied to Chilean pensions specifically, we found a few papers on using ML to model pension funds in other countries. A 2019 paper used Boosted Decision Tree Regression to predict the value of pensions for retirees in Peru[Aguirre *et al.* (2019)]. Another 2018 paper used a 2-layer neural network to optimize asset allocations for American target-based defined contribution pension plans [Li & Forsyth (2019)]. Both papers achieved promising results.

But these papers only predicted the future value of pension funds, not which pension fund plans retirees should opt for. A recent academic survey of ML in financial planning mentions that "This area of finance has been relatively immune to ML technology, except for a few exceptions such as high-frequency trading and credit scoring for loans" [Mulvey (2017)].

Recent non-ML work that has studied the retirement plan decisions of Chileans from a statistical perspective have been [Otero *et al.* (2019)] and [Illanes & Padi (2019)]. These two papers run linear regressions to estimate pension outcomes such as an introduced quality variable over the retiree's information such as sex, age, accumulated savings and beneficiaries. This research helped us to select which variables and information we should use.

## 3 Dataset Details & Data Pipeline

With the aim of predicting the pension option chosen by retirees, we used the data from the Chilean quotation platform SCOMP.[1] This electronic system is where individuals receive pension offers by life insurance companies. Participation is mandatory for those who have above a certain level of accumulated savings, thereby making it is a very rich public database that provides some personal information in an anonymous manner. An example of part of a row in the data can be seen below:

| **Retiree Sample** | **1** | **2** | **3** |
|---|---|---|---|
| Retiree Reason | Age (>=65) | Age (>=60) | Age (>=65) |
| Retiree Gender | Male | Female | Male |
| Birth Year | 1944 | 1957 | 1945 |
| Year of Pension Plan Acceptance | 2010 | 2018 | 2010 |
| Total Savings | 1242 UF | 2241 UF | 810 UF |
| Number of Dependents | 3 | 0 | 2 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| Disabled? | No | No | No |
| Worked Heavy Labor? | No | No | No |
| Community of Residence | Penco | Concepción | La Unión |
| Annual inflation Rate at Retirement | 1.41% | 2.43% | 1.41% |
| Poverty Rate of Community | 8.6% | 7.7% | 13.3% |
| **Pension Modality** | **Class 4** | **Class 2** | **Class 1** |

The data has been collected since August 2004, when SCOMP was created, until September 2020 as shown in Figure 2. The raw data contains over 8.6 million pension plan requests, 109.7 million retirement plan offers, and 0.68 million accepted pension plans by retirees. These datasets considered 100 descriptive columns including the retiree's saving amount, the number of dependents the retiree has, whether a financial advisor was used, and whether the retiree accepted the offer. Naturally, the dataset was in Spanish, but we were able to translate the relevant portions and clean it to make it

[1] Sistema de Consultas y Ofertas de Montos de Pensión

possible to train our models. Additionally, the size of the dataset presented some challenges, but with some additional pre-processing, we were able to merge the many individual files and filter the relevant data to construct our final dataset.

We want to focus only on the primary categories of retirees: early and old-age. So, we excluded those individuals that retired due to disability or survivorship (beneficiaries from a deceased retiree). Furthermore, important regulatory changes occurred in 2008, and so we consider retirees who accepted a pension plan offer between January 2009 and September 2020, leaving us with data on 400,000 retirees.

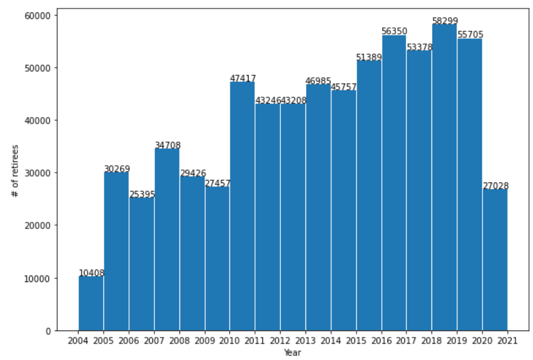


Figure 1: Total SCOMP retirees per year between August 2004 and September 2020.

Because we want to apply the model toward forecasting purposes in that we want to approximate a financial advisor's advice, we filtered out all ex-post variables so as to only have information that an advisor would have at the time of the decision. And further, we supplemented the existing dataset with demographic information linked to the retiree's location and economic data like interest rates, inflation, and unemployment levels over time to account for any external influences that may have altered the decisions recommended by advisors.

After this basic cleaning, we present the distribution of our output (Plan Modality):

1. Immediate annuity program: 123,865 cases.
2. Temporal programmed withdrawal with deferred annuity program: 137,875 cases.
3. Immediate annuity program with programmed withdrawal: 944 cases.
4. Programmed withdrawal: 144,071 cases.

Therefore, since Class 3 is almost never selected we decided to combine it with Class 4, hence creating a final subset of 405,811 samples that we used. The balance of the plan modality per year chosen by our final subset of retirees is displayed in Figure 2.

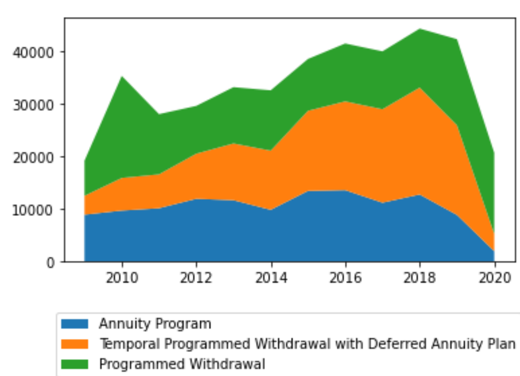


Figure 2: Selection of Plan Modality (our output) by our final subset of SCOMP retirees per year.

# 4 Methodology

As our model seeks to simulate the options produced by financial advisors, we assumed that the pension plans chosen by each retiree were the optimal choice, and that thus both the human error and the Bayes error was 0. Although we recognized that financial planners might sometimes make suboptimal decisions for retirees, in the absence of definite proof as to which pension plan was perfect (which in the field of financial planning is impossible), we chose to treat financial planner decisions as the gold standard.

For the implementation, we chose to use the Scikit-learn framework. We began our analysis by running a baseline model using logistic regression on the un-normalized dataset. Next, we produced two more neural network models, the first with one hidden layer with 20 neurons and the second with two hidden layers with 20 neurons each as well. In order to reduce potential over-fitting, we then tried adding regularization to both neural network models as well.

Our accuracy was still not high, so we supplemented our dataset with demographic and socioeconomic data. Additionally, we noticed we had a lot of features due to the one-hot encoding of many categorical variables. Of note here was the variable of commune, which had more than 350 options. In order to reduce the number of features present in our dataset, we tried running K-means clustering on the commune data to see if we could reduce the number of features generated by one-hot encoding into a smaller number of features generated by the K-means clusters. This was important because a number of communes had very limited data, and it was possible this was skeing the results, so with some analysis of the data distribution, particularly how many communes had very little data, and some tuning, 20 clusters were chosen. We one-hot encoded this new location cluster feature and then merged it with the demographic data associated with each cluster's communes. This resulted in a total of 38 features, compared to the initial nearly 400. However, this attempt produced only marginally better results, increasing our test accuracy from 63% to 64%. But it did provide a slight regularization effect.

Unsatisfied with these results, we then tried a really large neural network with five layers of 200 neurons each. This model ended up overfitting the data, producing 75% accuracy on the training dataset and 61% accuracy on the test dataset. Repeated with K-means clustering, we had some regularization but again little increase in overall test accuracy. In order to create more of a regularization effect, we attempted other techniques, including early stopping and dropout. Another model architecture tried included skip connections in order to address potential training instability, but regardless, all of these additions did little to improve our overall model performance compared to the initial accuracy levels.

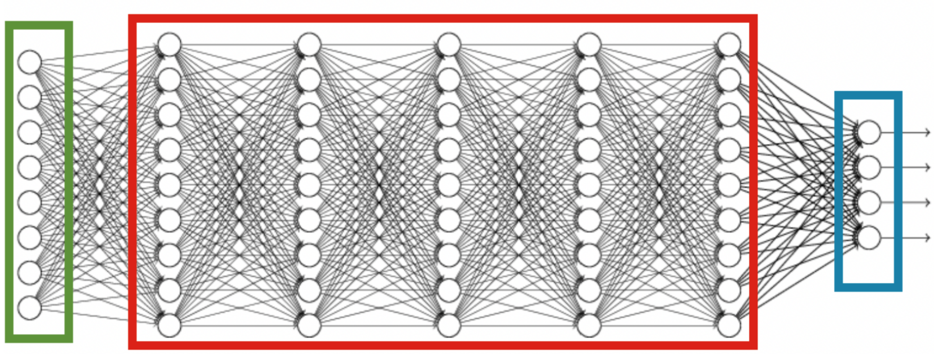

Figure 3: A compressed model of the final neural network architecture we tried. Nodes in green indicate the input fields, nodes in red indicate the five hidden layers, and the final output layer includes softmax values indicating the probability each retiree selected X pension plan option.

## 5 Results

| Model | Validation Accuracy | Validation Loss | Test Accuracy |
|---|---|---|---|
| Logistic Regression on Unenhanced Data | 0.352 | N/A | 0.354 |
| Logistic Regression on Enhanced Data | 0.605 | N/A | 0.602 |
| 2-Layer Neural Network w/o K-Means | 0.645 | 0.760 | 0.633 |
| 5-Layer Neural Network w/o K-Means | 0.644 | 0.725 | 0.628 |
| 5-Layer Neural Network w/ K-Means | 0.655 | 0.738 | 0.631 |

Despite our best efforts and the variety of models we trained, we were unable to generate results that reached past 63.3% accuracy on our test datasets. Adding more layers and greatly increasing the number of neurons per layer helped slightly increasing training accuracy but ultimately did not change test accuracy much, and thus only produced overfitting. We concluded that the lower than anticipated test accuracy likely indicated that our model still had a lot of avoidable bias that likely arose from deficiencies in our dataset. Though we spent a great deal of time during our project finding more demographic and socioeconomic data and adding that to our dataset, this did not improve accuracy drastically. So, it became clear that without access to individual retirees' personal preferences regarding things like risk tolerance or retirement goals, it is difficult to attain a higher level of predictive accuracy.

| | Precision | Recall | F1 Score | Support |
|---|---|---|---|---|
| Class 1 | 0.55 | 0.42 | 0.48 | 3899 |
| Class 2 | 0.59 | 0.60 | 0.59 | 3900 |
| Class 4 | 0.66 | 0.80 | 0.72 | 4353 |

Our total accuracy of 64% test still marks a remarkable improvement over both a naive Bayes classifier and our baseline logistic model. Had our model just randomly assigned pension plans to retirees, it would have only produced a test accuracy of at best about 37% (the proportion of the highest chosen plan). This means that our model could be used by government officials to analyze overall trends in retiree pension plan decisions without access to individual preference data.

## 6 Conclusion

Overall, we managed to produce classifiers that were able to make correct pension plan predictions for a large majority (nearly 2/3) of the retirees. Although we were unable to reach the very high accuracy we had originally hoped for in our bid to replace the services of financial planners, the fact that we were able to generate results with as good accuracy as we produced with just each pensioner's socioeconomic, demographic, and basic financial data proves extremely promising for further research in this subject. Compared to the hard sciences, this accuracy is low, but for social sciences, it's actually extremely high since it's very difficult to explain or predict human behavior even with a lot of data.

As most of our model's faults seemed to arise from a lack of relevant information concerning each pensioner's risk tolerance and financial goals (important variables that factor into a retiree's pension decision that we simply did not have any data on or could not predict from the data given), future actions we may try can include surveying retirees as to their financial risk tolerance and retirement goals, and then producing more accurate models given that information.

## 7 Contributions

All three members of the team were heavily involved in each aspect of the project, but certain people focused more in specific areas based on their skills. Sebastián Espinoza, being from Chile, took the lead on understanding and translating the industry context, acquiring and cleaning the data, which included two clarifying meetings with government officials, running basic models, and helping in the advanced models. Akshay Malhotra, as a Finance PhD student, took the lead on the academic literature research, conducting advanced cleaning of the data and developing the initial models along with regularizing and applying KMeans to the latter models. Daniel Zhang, as a Computer

Science student, took the lead on comparing different deep learning strategies, such as merging external datasets, neural network architectures, dealing with the bias-variance trade-off and running the advanced models.

## 8 Bibliography

## References


Aguirre, Jimmy Armas, Ladera, Jhonathan Espinoza, Castillo, Brian Dueñas, & Mayorga, Santiago Aguirre. 2019. Predictive analysis for calculating the valuation of the affiliated fund of a private pension system using machine learning techniques and tools. *LACCEI2019*.

Dube, Ryan. 2016. Chileans Protest Against Private Pensions. *Wall Street Journal*, 8.

Illanes, Gastón, & Padi, Manisha. 2019. Retirement Policy and Annuity Market Equilibria: Evidence from Chile. *Working paper*, 9.

Li, Yuying, & Forsyth, Peter. 2019. A Data Driven Neural Network Approach to Optimal Asset Allocation for Target Based Defined Contribution Pension Plans. *Insurance: Mathematics and Economics*, **86**.

Mulvey, John M. 2017. Machine Learning and Financial Planning. *IEEE Potentials*, **36**.

Otero, Andrés, Palominos, Claudio, Quintanilla, Ximena, & Céspedes, Vicente. 2019. Chilean Annuity Market: Annuity choice in a segmented pension market. *Working paper*, 5.