Data 610-9080 Decision Management Systems

Assignment 3 Decision Tree Model Development using Cognos Analytics

Prepared by Kierra Young

Email: [kyoung93@student.umgc.edu](mailto:kyoung93@student.umgc.edu)

Prepared for Professor Laila Moretto

11/10/2020

**Introduction**

The dataset that is being used for this report is the billionaire.csv. The contents of this data include 2614 cases of people who have became billionaires in the past a couple hundred years. There are twenty-two different variables in this dataset. The variables describe the characteristics each of the billionaires possess and reveals what factors contributed to these people’s success over time. The types of variables that were chosen for this report were ones that would lead to answering the two following questions: “What age are billionaires likely to become billionaires?” and “Do wealth types contribute to one becoming a billionaire?” After the data was cleansed and null values were removed it was simpler to observe what variables could answer these questions. An image of the visible variables used in this report is in the Appendix as Figure 1. Visible Variables.

**Predictive Models Analysis**

There were two predictive models made. One model focused on the outcome of the age of billionaires while the other focused on the wealth type. The goal was to investigate the variables and see if there was any insight between the predictions made and the questions asked.

Age of Billionaires Decision Tree Model

To begin, the first predictive model that was made was to indicate what age most people in the past couple hundred years have become billionaires. According to the article “Forbes 2017 Billionaires List: Meet the Richest People on The Planet”, it says the following:

“Mainland China had the newest ten-figure fortunes with 76. The U.S. was second with 25. Notable newbies include Patagonia sportswear founder Yvon Chouinard; Viking Cruises founder Torstein Hagen of Norway; U.S. hedge fund tycoon Cliff Asness and two of his partners; and John and Patrick Collison, Irish citizens who cofounded San Francisco-based Stripe, which enables online payments. John Collison, age 26, is now the world’s youngest self-made billionaire, just two months younger than Snapchat’s Evan Spiegel. The Collisons are two of just four self-made billionaires in their 20s (the other two are the Snapchat cofounders). There are 56 billionaires under age 40, down from 66 last year, after some aged out and others dropped below the $1 billion mark.”( Chen, 2018)

From this article it can be said that people are becoming billionaires at a much younger age. The predictive model in Figure 2. Age of Billionaires illustrates how long it took previous billionaires to become billionaires and what variables contributed to it. To be specific the exact variable being predicted is the age. This variable is important to predict because it allows people to know what age they might be when they become a billionaire depending on certain factors. The variables that are being used to predict the outcome of age are founded, gender, year, worth in billions, GDP, industry, rank, was political, relationship, wealth type, region, citizenship, and category.

Age of Billionaire Results/Analysis

The results for the decision model tree made for predicting the age of billionaires indicate that the relationship between the input variables and target variable chosen has an insight strength of 41.1 percent. The decision model details say that the exact variables founded, relationship, country code, and gpd predict age with a predictive strength of 41.5%. In addition, the year is the least important predictor of age. Year only improves the predictive strength by about <0.1%. Therefore, the strongest variables are founded, relationship, country code, and gpd. The single best predictor variable is founded. If a company was founded in <1975 the predictor strength is 59%. The combination of variables that seem to give the best answer are founded, gpd, relationship, country code. These are the main variables that drive the prediction without them the prediction would not be as accurate. For example, when replacing the variable country code with industry the prediction percentage went down. The combination above gives the most accurate predication.

Age of Billionaire Analysis of Rules/ Sunburst Diagram

There are many rules in the age of billionaire’s decision tree models. The top five best rules according to the decision tree model are as follows: first, the variable founded is < 1959 and <1975 by a person with a career status of either a chairman, CEO, president, founder/chairman, founder/CEO, or employee that have a GDP less than 3.79E11 and a country code equal to USA, CAN, CHE, FRA, PHL, SWE, AUS, SGP, ITA, ZAF, Taiwan, BEL, ISR, EGY, MAR, AUT then they are predicted to have become billionaires at age 83.33. Second when the variable founded is 1959 ≤ founded < 1975 by a person with the career status of relation, chairman, CEO, president, relation and chairman, founder/chairman, founder/CEO, employee with a gdp < 3.79E11 and country code = USA, CAN, CHE, FRA, PHL, SWE, AUS, SGP, ITA, ZAF, Taiwan, BEL, ISR, EGY, MAR, AUT then they are predicted to have become billionaires at 76.03. Third, when variable founded is founded < 1975 with a career status equal to relation, chairman, CEO, president, relation and chairman, founder/chairman, founder/CEO, employee with a gdp < 3.79E11 and a country code = BRA, DEU, JPN, MEX, CHL, MYS, HKG, GBR, THA, ESP, IRL, COL, IND, TUR, IDN, SAU, NOR, RUS, KAZ, CHN, PER, POL, UKR then they are predicated to be billionaires at 72.32. Fourth, when variable founded is founded < 1975 and founded <1930 and the relationship = founder, owner, investor, Chairman and Chief Executive Officer, former CEO, founder and chairman, Chairman, founder and CEO, founder/relation, former chairman and CEO with a country code = USA, DEU, CHL, MYS, HKG, GBR, THA, ESP, FRA, SWE, COL, AUS, IND, TUR, ZAF, IDN, Taiwan, KOR, ISR, NOR, EGY, DEN, and gdp < 3.68E9 it is predicted that a billionaire would be 71.79 in age. Lastly, when the variable founded < 1975 with a career status of relation, chairman, CEO, president, relation and chairman, founder/chairman, founder/CEO, employee with a gdp ≥ 3.79E11 in the industry of either Retail, Restaurant, Technology-Medical, Real Estate, Mining and metals, Energy, Consumer, Hedge funds, Diversified financial, Venture Capital they are predicted to be a billionaire at 69.24 in age. Likewise, Figure 3. Age of Billionaires Sunburst Diagram gives knowledge on the top 5 age values and it pinpoints where all the values are in the diagram. It shows the founder surrounding the age in the middle because age is the target variable. It also shows the standard deviations of the variables, total number of records, percentages, and other appropriate information on the variables.

Rules Applying to My Organization

Furthermore, the rules found are very much useful for my organization which is the data analytics industry. These rules are useful because they give an average of the predicted value, a detailed description of the rules, and it also shows the records/ percentages to prove that the rules are accurate. I would implement the use of the rule-base approach into machine learning AI. The approach would be able to manipulate information and use it in a beneficial way for future data related projects. The models would be accepted as information or the basis for making data related predictions to the company can make smart decisions when needed an exact average of number such as age, time, day, or etc.

Wealth Type of Billionaires Decision Tree Model

To continue, the second predictive model was made to discover if the variable, wealth type, contributes to people becoming billionaires. There was an article called “Wealth of Nations or Wealth of Persons: World Billionaires and Sector Concentration”, that examined whether wealth in the 21st century is “closely related to the sector concentration in billionaire-generating countries.”( YAKIŞIK, 2013 ) The following gives a snippet of the article:

“By the help of the technological advances and the speedy increasing in deregulation in financial sector, this huge increase in the number of billionaires and wealth had never been created during the last decades. After the industrial revolution, parallel to the industrialization process, world billionaires emerged in early industrialized countries.” (YAKIŞIK, 2013)

This article explains how wealth has increased over the years due to advancements in technology. The decision tree will see if the following variables: year, gender, region, country code, worth billions, citizenship, age, founded, GPD, industry, rank, was political, name, category, company type, and relationship can predict the outcome of the target variable wealth. This target variable is important to predict because inside the variable it has the main labels of the type wealth each person possess in the data. Figure 4. Wealth Type Tree Diagram is in the Appendix.

Wealth Type of Billionaires Results/Analysis

The results from this decision tree for predicting the outcome of the variable wealth type indicate that the relationship between the variables mentioned above and wealth type have a predictive insight of 89.7%. The variables that are the strongest predictors are relationship and category which predict wealth type by 89.7%. The single best predictor variable is relationship. The relationship strength is 100% the dictator of what the wealth type outcome will be. The combination of variables that seems to give the best answer are the following: year, gender, region, country code, worth billions, citizenship, age, founded, GPD, industry, rank, was political, name, category, company type, and relationship. These were all the variables used to predict wealth type outcome and they are the best because it gives a variety of knowledge to the dataset making it more accurate.

Wealth Type of a Billionaire Analysis of Rules/ Sunburst Diagram

There are many rules in the decision tree diagram for wealth type. The top five rules are as follows: first wealth type if predicted at 100% when relationship = owner, chairman, Chairman and Chief Executive Officer, CEO, founder and CEO, employee, and founder/relation with a category = Resource Related. Second wealth type is predicted at 97% when relationship = relation, president, Chairman, founder/CEO, and executive chairman with a category = Resource Related, energy. Third wealth type is predicted at 96% when relationship = relation, president, Chairman, founder/CEO, and executive chairman with a category = Non-Traded Sectors, New Sectors, Traded Sectors. Fourth wealth type is predicted at 96% when relationship = founder, former CEO. Lastly wealth type is predicted at 95% when relationship = relation, president, Chairman, founder/CEO, and executive chairman with category = Financial. Likewise, the sunburst diagram breaks down the variable wealth type more clearly and by colors. However, it shows the top five categories in the variable wealth type that are most successful. The relationship variable is surrounding the wealth type proving that the relationship variable is the greatest dictator of the outcome for wealth type. Executive is a category colored in blue and has 734 records of billionaires who were founders, CEOs, and relationships. The green color are people who achieved billionaire status through founder-non finance and there are 599 records. Yellow are people who inherited their billionaire status and there are 217 records. Then there is orange in the color wheel that is for billionaires in privatized and sector with a record of 439. Then there is a purple color for self-made financial billionaires with a record of 157. Figure 5. Wealth Type Sunburst diagram is in the Appendix.

Rules Applying to My Organization

The use of the rules found in the decision tree for wealth type of a billionaire would be useful in my organization. The data industry is always analyzing for business, investors, entrepreneurs, and other ways for people to make money. In an article it was discussing a young 39 old man who became a billion based off a certain wealth type and strategy he had. The following is a quote from the article:

“Wang, 39, is the latest in a recent rush of newly-minted Chinese billionaires as **founders** take their startups public. Five tech IPOs have helped at least 16 people ring up a combined $48 billion of net worth at the time of the IPOs, according to data compiled by the Bloomberg Billionaires Index. That includes **founders** of smartphone maker Xiaomi Corp. and Colin Huang, who is now worth $11.9 billion after the successful listing of his e-commerce platform Pinduoduo Inc.”( Chen, 2018)

People such as Wang have financial advisors and data analyst to make sure that the investments, he makes will profit one day. I would implement the use of the rule-based approach by looking at the wealth type he had and seeing if the relationship strength is strong in that category field the client is in. Wang is a founder and has a technology company. The model would be accepted by displaying how people that are founders/ CEOs in the tech and financial industry are more likely to be billionaires.

Additional Models & Conclusion

The model Figure 6. Age of Billionaire Comparison is a comparison of the original tree model to a second one. I used fewer variables and still achieved the same predicted outcome, but the fewer variable method has greater insight now at 41.3% shown as Chart B. Likewise Figure 7 Age of Billionaire Additional Model is a more detailed in the form of a heat map. It shows that for age, founder and relation are the most important categories of relationship with a total value of 112,534 (83.9 % of the total). Likewise Figure 8. Wealth Type Comparison is the original tree diagram to a new tree diagram showing the rules. The original used all the drivers, but the side chart A only used wealth type as the target and relationship and category as the input variables. The results stayed the same. Furthermore, Figure 9. Wealth Type Additional Model is a column model for founded by gender and wealth type. This model is still divided by colors like the sunburst diagram, but it is also broken down into gender which gives more insights; male is the most frequently occurring category of gender with a count of 1,941 items (90.4 % of the total), while female is the other 9.6% of billionaires.

References

Chen, L. (2018, September 18). 39-Year-Old Founder's Wealth Jumps to $5.3 Billion After IPO.

https://www.bloomberg.com/news/articles/2018-09-20/meituan-rises-on-debut-in-hong-kong-after-4-2-billion-ipo

Dolan, Kerry A. (2017, Mar 19) *Forbes 2017 Billionaires List: Meet The Richest People On The*

*Planet*. www.forbes.com/sites/kerryadolan/2017/03/20/forbes-2017-billionaires-list-meet-the-richest-people-on-the-planet/.

YAKIŞIK, H. (2013, June 1). *Wealth of Nations or Wealth of Persons: World Billionaires* ...,

from http://websitem.karatekin.edu.tr/user\_files/haruny/files/karatekin-makale.pdf

Appendix

Figure 1. Visible Variables for Billionaires

Graphical user interface, table

Description automatically generated

Figure 2. Billionaire Age Tree Diagram Average Ages

Chart

Description automatically generated

Figure 3. Billionaire Sunburst Tree Diagram

Chart, sunburst chart

Description automatically generated

Figure 4. Wealth Type Tree Diagram

Graphical user interface

Description automatically generated

Figure 5. Wealth Type Sunburst Diagram

Chart, sunburst chart

Description automatically generated

Figure 6. Age of Billionaire Comparison on Sunburst Tree Diagram

Chart, sunburst chart

Description automatically generated

Figure 7. Age of Billionaire Additional Model

Graphical user interface, application, Teams

Description automatically generated

Figure 8. Wealth Type Comparison Rules Tree

Graphical user interface, application, Teams

Description automatically generated

Figure 9. Wealth Type Additional Model

Chart, bar chart

Description automatically generated