- ORIGINAL ARTICLE -

A Case Study of Forecasting Elections Results: Beyond Prediction based on Business Intelligence

Un Caso de Estudio de Pronóstico de los Resultados Electorales: Más allá de la Predicción Basada en Business Intelligence

Antonio Lorenzo[1,2], José A. Olivas[1]

[1] *SMILe (**S**oft **M**anagement of **I**nternet and **L**earning).*
Information Technologies and Systems Institute,
University of Castilla La Mancha, Ciudad Real, Spain.
JoseAngel.Olivas@uclm.es
[2] *Coordinator of the Department of Business Intelligence,*
Castilla La Mancha Government, Toledo, Spain
alorenzo@jccm.es

Abstract

In the field of data analysis, it is common not to distinguish clearly between prediction and forecast. Although the results of both processes may tend to converge, the mechanisms used in each case tend to be completely different. Prediction has to do with statistical extrapolation and estimation and forecasting can consider expert judgments on the subject. A methodology is proposed to carry out this latter task, with a mechanism that uses both historical and current data with the judgement of an expert. The methodology is applied to the case study of the Spanish general elections of April 2019.

Keywords: Business Intelligence, Expert Knowledge, Forecast, Methodology, Prediction.

Resumen

En el campo del análisis de datos, es común no distinguir claramente entre predicción y pronóstico. Aunque los resultados de ambos procesos pueden tender a converger, los mecanismos utilizados en cada caso tienden a ser completamente diferentes. La predicción tiene que ver con la extrapolación estadística y la estimación y el pronóstico puede considerar el juicio de expertos sobre el tema. Se propone una metodología para llevar a cabo esta última tarea, con un mecanismo que utiliza tanto datos históricos como actuales con el juicio de un experto para afinar el resultado. La metodología se aplica al estudio de caso de las elecciones generales españolas de abril de 2019.

Palabras claves: Conocimiento experto, Inteligencia de negocio, Predicción, Pronóstico.

1. Introduction

Usually prediction and forecasting are terms that are used interchangeably to venture future events. The prediction uses historical or current data from time series, which statistically analyzed through the application of a mathematical model, estimates what will happen in the future. Normally, the prediction has a high percentage of successes since it is possible to determine what the weather will be tomorrow by knowing the weather of today and the evolution of temperature, wind speed or atmospheric pressure. On the other hand, it is extremely complex to predict the weather within a year, since the evolution of weather from today to within a year provides many different scenarios causing the success rate to be very low. Tomorrow's weather is determined by today's weather and the weather of the day after tomorrow, tomorrow's... and so on. In the prediction, both events (past and future) can be related, that is, there is a connection and future data can be deduced, to a greater or lesser extent, from historical and current data. The prediction moves in situations of certainty regarding the fact to predict, using historical information to project it into the future following the trend with a high probability of success. For example, the population of Toledo (Spain), since 1998, has progressively evolved: from 1998 with 66,989 inhabitants to 2017 with 83,741 inhabitants. It is relatively easy to predict the population in 2020 will be around 84,000 but it is difficult to determine the population in Toledo in 2030 because the evolution may be affected by other factors that make it not follow a linear growth.

Forecasting is much more complex and difficult to carry out since the future must be estimated in situations of uncertainty. The forecast makes it possible to anticipate a specific event based on a small set of alternatives, for example, in football

(soccer) betting the specific event is "to know if a team is going to win the match" and the alternatives are limited: win, draw or lose (1 X 2). In the stock market, the specific event is to know "If there is going to be a crack in the stock market that makes it collapse". For example, the vote in favor of Brexit caused a 12.35% drop in the Ibex35 index (Spain) in a single day. In forecasting there is a small set of differentiating facts that cause a turning point in the trend that makes the result vary substantially. The forecasting is the inflexion point in the trend that causes the results to change substantially with respect to what was expected.

BI tools perform a descriptive analysis of the data and help us understand *what happened* and *why it happened*. The traditional BI tools are called Operational Business Intelligence (OBI). But there is another type of Business Intelligence, called Analytic Business Intelligence (ABI), that wants to predict what will happen in the future with respect to a specific event, that is why it uses predictive analysis techniques (extracts knowledge in the form of patterns, models or trends) and prescriptive analysis techniques (what we should have done for a certain event to occur). Operational BI tools are not enough to make forecasts; a new method is proposed using experts to add additional information to the model to improve prediction. As an application of the proposed method, a case study will be carried out in the Spanish Elections to forecast the results in 2019.

The application of AI and statistical methods to electoral forecasting has been widely studied by the scientific community. Stoetzer [1] offer a dynamic Bayesian forecasting model for multiparty elections. It combines data from published pre-election public opinion polls with information from fundamentals-based forecasting models. Graefe [2] test the effects of forecasting accuracy, we applied three evidence-based guidelines to 19 published regression models used for forecasting 154 elections in Australia, Canada, Italy, Japan, Netherlands, Portugal, Spain, Turkey, U.K., and the U.S. Huber [3], forecasts the results of the German State Elections based on polls data from different institutes. To predict the participation of a single vote in multiparty elections, the range of methods varies from basic methods such as averaging over methods based on nonparametric regression to dynamic linear models. Graefe [4], this study analyzes the relative accuracy of experts, polls, and the so-called 'fundamentals' in predicting the popular vote in the four U.S. presidential elections from 2004 to 2016. Combining expert forecasts and polls with a fundamentals-based reference class forecasting reduced the error of experts and polls by 24% and 19%, respectively. Kaufman [5], though used frequently in machine learning, boosted decision trees are largely unused in political science, despite many useful properties. This article explains how to use one variant of boosted decision trees, AdaBoosted decision trees (ADTs), for social science predictions, in particular, predicting county-level vote shares in U.S. presidential elections. De Rosa [6] intends to contribute to the debate via a critical analysis of the way big data is used to predict election outcomes. The focus of the analysis is the first round of the 2017 French presidential elections, and the voting forecasts produced during the campaign. Normally in electoral forecasting models based on the mathematical analysis of the data obtained from the polls, the judgment of an expert is not applied.

Another way to make predictions are dashboards, through historical data analysis. Normally, Operational BI tools have some functionality that allows, varying the initial conditions, to see how they influence the final result. It's a simple rule of three. For example, if we assume that a loaf of bread weighs 250kgr, with 100 kgr of flour I produce 2,300 loaves, and if I use 150 kgr of flour I will produce 1,854 loaves of bread. But you can make a more complex prediction, for example, you can predict what would happen if we want to increase last year's sales by 10%. This result in a compound change, it can happen that production costs are expected to rise by 15% and margins are expected to fall by 5% as a result of rising costs. Or on the contrary, production costs only increase by 5% and the margin increases by up to 15%.

2. Goal

In the 2019 General Elections in Spain, the left wing is formed by the *Partido Socialista Obrero Español (PSOE)* and *Unidas Podemos (UP)*. The right wing is formed by the *Partido Popular (PP)*, *Ciudadanos* (Cs) and *Vox*.

In Spain the Official Organism in charge of carrying out electoral opinion studies (polls) is the Sociological Investigations Centre[1] (CIS). For the 2019 General Elections the CIS made the following forecast (Table 1):

Table 1. General Elections of Spain. CIS' forecasting.

	Members of Parliament General Elections 2016.	**CIS' forecast General Elections 2019. Members of Parliament.**
Left-wing (PSOE+UP)	156	180
Right-wing (PP+Cs+Vox)	169	135

[1] www.cis.es

The aim of this article is to define a methodology and apply it to the results of the 2019 General Elections in Spain to improve the forecasts made by CIS on the number of Members of Parliament obtained by the left wing (*PSOE+UP*) and by the right wing (*PP+Cs+Vox*).

3. Methodology to make forecasting proposal.

In order to establish a hypothesis that leads us to make a forecast, it is necessary to have the knowledge of an expert, someone who, even if it is *a posteriori*, can determine which differentiating facts have defined the final result. In the case of football, most fans know *a posteriori* why a football team has won or lost the match. The forecasting is easy to explain *a posteriori*, because a descriptive analysis is made to know what and why it happened, but it is difficult to establish *a priori* because specific facts that determine the result are unknown.

Prediction is the extrapolation of current data into the future. Forecasting is the turning point that breaks the trend. For example, in the semifinals of the European Football Cup in 2018 and 2019, in the first match Barcelona (FCB) won 3-0, both Rome in 2018 and Liverpool in 2019. In the second match, Barcelona FCB lost the semi-final to both Roma and Liverpool. The prediction would have said that Barcelona would have won semifinals because it had a considerable advantage in the first match, but in the second match, the trend broke and lost. For this, in any forecasting model, there must be knowledge of an expert, who applying the correct hypotheses, determines an accurate forecasting. For example, in the first match, despite winning the Barcelona showed a worse game than Liverpool. Some experts predicted that Barcelona could lose the second match because Liverpool had a much better game than Barcelona.

Traditionally, statistical techniques have been used to make predictions. These include time series, the ARIMA algorithm or causal models based on statistical regression. In recent years, the use of soft computing techniques (mainly neural networks) has begun. For example, Atsalakis et al. [8] [9] carries out a comparative study between different statistical techniques and neural networks to try to predict stock market values, and more recently, Vaidehi et al. [7] has joined machine learning algorithms with fuzzy logic obtaining better results.

Soft computing techniques improve the effectiveness of forecasting models with respect to traditional statistical techniques, but they are still not sufficient because they work well when the case to be forecast has been previously trained by the system, but when a new case is presented, which is not the type that has been used to train the model, soft computing techniques have a low success rate.

The main idea is to define a methodology to improve the results of statistical and soft computing techniques based on the analysis of historical data from a dashboard and adding the opinion of an expert who treats it with fuzzy logic. A methodology is proposed to obtain a model that generates a cognitive system, that from the historical data, plus the knowledge of an expert that determines which exogenous factors have influence the final result. This methodology consists of the following phases (Fig.1):

1) A dashboard is made with the historical data of the event to be predicted, carrying out a descriptive analysis of the data. For this purpose, market-leading tools such as *Qlikview*, *PowerBI* or *Tableu* are used. The functionalities of these tools are used to extrapolate the result to the future, observing the result. In the case of *Qlikview*, the functionality is called "*What-if analisys*".
2) Depending on the event to forecast, which can be from different sources (radio, press, television, internet...), an expert should be established which factors may have an influence on the forecast. For each of these characteristics, the range of values (linguistic labels) over which it can vary will be determined, e.g. "a lot", "a little", "not at all".
3) The factors indicated by the expert are applied to each of the elements to be forecast, establishing their influence on each representative element. It can happen that the expert also indicates the conditions to take into account.
4) For each one of the representative elements of the event to forecast, an extended forecast is established, where together with the historical data, the conditioning factors and each one of the characteristics that the expert has determined, a new forecast is made.

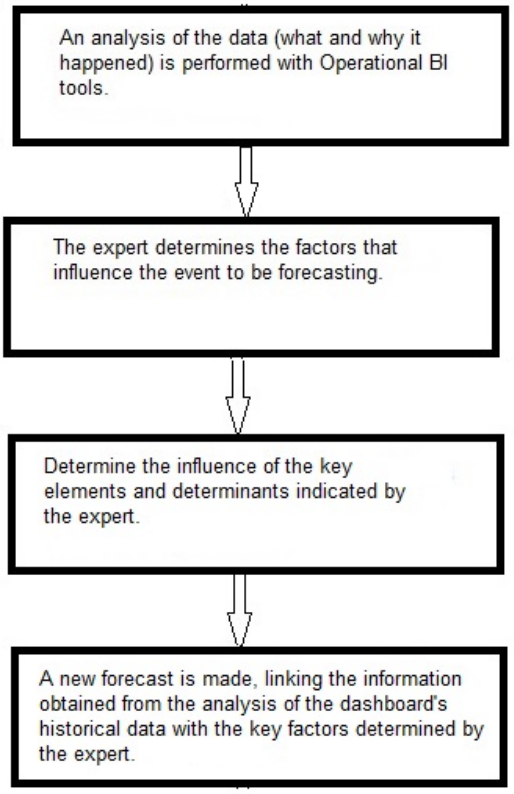

Fig. 1. Phases of the forecasting methodology.

4. Case study: General Elections in Spain 2019.

4.1. Elaboration of a dashboard with historical information: Operational Data Analysis.

Historical electoral data can be analyzed with BI tools in order to extract knowledge. BI tools perform a descriptive analysis of the data and help us understand *what happened* and *why it happened*.

A dashboard has been created using the BI *Qlikview* tool with the electoral results for the 2004, 2008, 2011, 2015 and 2016 General Elections (Fig. 2). The *Qlikview* tool has been selected for the dashboard because it is one of the leading tools in the market according to the Gartner 2019 Magic Quadrant [10], along with *PowerBI* and *Tableau*. *Qlikview* uses Associative Logic (AQL) technology, unlike multidimensional data (Pentaho, Oracle Business Intelligence, Microsoft Analysis Services...), there is no hierarchical structure between the data, but there is a many to many relationship between the facts so that you can make a selection without having previously selected the parent data in the hierarchy. For example, with multidimensional data analysis tools a hierarchy is established in such a way that in order to select a city, the country and province must first be selected. The *Associative Query Logic (AQL)* allows selecting the municipality and automatically those country and province data compatible with the selected municipality are available. This technology allows searching and interacts with the data in the same way they are thought, in an associative way. Users can instantly see connections and relationships between data from different sources.

From the descriptive analysis of the data, from 2004 to 2016, important information can be extracted from the provinces where the right-wing has always won, or to date, has a high probability of winning (Cantabria, Ávila, Palencia, Segovia, Valladolid, Zamora, Cuenca, Ceuta, Alicante, Castellón, Madrid, Melilla and Murcia, Málaga, Zaragoza, Salamanca, Ciudad Real, Guadalajara, Valencia, La Coruña, Lugo, Ourense, Pontevedra and La Rioja) and provinces with a high probability that the left-wing will win (Cádiz, Huelva, Jaén, Sevilla, Huesca, Barcelona, Lleida, Tarragona and Álava).

Qlikview make predictions through the "*What-if analysis*" functionality. This technique, which allows changes to be evaluated before they occur, helps predict the outcome by reducing risk, establishing different scenarios by varying the original data and seeing their impact on the final outcome. In Spain the form the function that establishes the relationship between the number of votes and the number of Members of Parliament obtained is carried out by means of the rule of D'Hondt [11]. The D'Hondt rule is a higher average method for allocating seats in proportional representation systems by electoral lists. The higher average methods are characterised by dividing through different divisors the totals of votes obtained by the different parties, producing sequences of decreasing quotients for each party and allocating the seats to the higher averages. It was created by the Belgian jurist Victor D'Hondt in 1878. The "*What if*" functionality can be used to vary the number of votes obtained by each political party in each province.

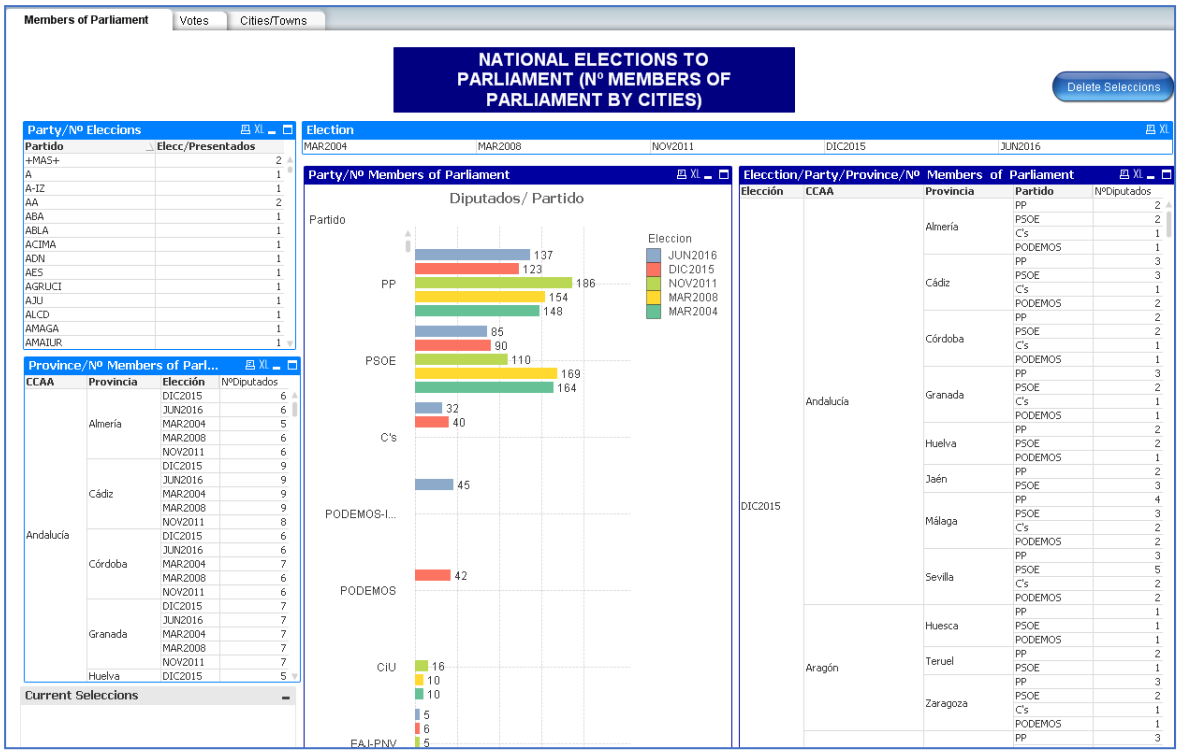

Fig. 2. Qlikview Dashboard for electoral results for the 2004, 2008, 2011, 2015 and 2016 General Elections in Spain.

In this way you can calculate, following the trend, the number of Members of Parliament in different simulation scenarios. Madrid has 36 Members of Parliament assigned to it in the 2016 General Elections. The following table 3 shows the real votes obtained by each political party in the 2016 General Elections in the province of Madrid and the simulation if PSOE, instead of obtaining 678,340 votes and 7 Members of Parliament would have obtained 950,000 votes. In the new distribution of Members of Parliament, the *PSOE* would have obtained 10 Members of Parliament at the expense of the *Partido Popular* and *Unidas Podemos*, who lose 2 and 1 Members of Parliament, respectively (Table 2).

Table 2. Results of General Elections 2016 in Madrid and simulation.

	Real Votes	Real Members of Parliament	Simulation Votes	Simulation Members of Parliament
Partido Popular (right-wing)	1,325,665	15	1,325,665	13
Unidas Podemos (left-wing)	737,885	8	737,885	7
PSOE (left-wing)	*678,340*	*7*	*950,000*	*10*
Ciudadanos (right-wing)	616,503	6	616,503	6

In the 2011, 2015 and 2016 General Elections, the right wing won. In the General Elections, in Spain, normally, when the left wing wins the turnout is higher than 70% and when the right wing wins it is lower. If the trend described in the dashboard data were followed, the 2019 General Elections would be won by the right wing and the turnout would be less than 70%. *In the 2019 General Elections, will the percentages of Members of Parliament obtained by each political party in each type of province be maintained?.Obviously not*, despite the power of *Qlikview*, this type of analysis is not enough to make forecasts because it is a projection on historical and current data.

4.2. Expert opinion.

In every Election, there are a few facts that are decisive in the final result and each Election is different from the other. After having consulted different sources of expert information it is determined that the elements that may have influence on the final result are:

1) Right wing's Fragmentation. For the first time since the 1977 democratic elections there will be five political parties with a voting intention of more than 10%, and of those, three are located in right-wing ideology (*Partido Popular*, *Ciudadanos* and *Vox*), and consequently the vote for this block will be fragmented.
2) Catalonia's independence is going to be key in the provinces of Barcelona and Tarragona. *Ciudadanos* (right-wing) is the

political party that won the 2017 elections to the Parliament of Catalonia for its defence of the Spanish nation against Catalan independence. Due to the opposition of *Vox* (right-wing) against Catalonia's independence, and the lukewarmness of the rest of the right-wing parties with Catalonia's independence, there will be a transfer of votes from the *Partido Popular* (right-wing) and *Ciudadanos* (right-wing) to *Vox* (right-wing).

3) The defence of the rural world and empty Spain. 5% of the Spanish population lives in 50% of the territory. It mainly affects the centre of Spain (Castilla y León, part of Castilla-La Mancha, part of Aragón, part of Extremadura, part of Galicia...), except Madrid. For years they have demanded road infrastructures, health and educational services, high-speed Internet networks... which have not been attended to. These are areas that have been depopulated due to the lack of attention from Public Administrations. *Ciudadanos* (right-wing) y *Vox* (right-wing) are the parties that has stood out most for the defence of the rural and traditional world (bulls, hunting, popular festivities, Internet access in rural areas...).
4) Disenchantment of the *Partido Popular* voter. The right-wing voter believes that his or her vote will be more effective if he or she votes for Vox, since the *Partido Popular* have had experience of government or the ability to influence it and have not met the expectations of many of them. The Popular Party has had numerous cases of corruption in the Comunidad of Madrid and in the Comunidad of Valencia.
5) Inmigración. With regard to immigration, it will be a factor that will have an influence in large provinces where a significant percentage of the population is immigrant [12]: Murcia, Malaga, Santa Cruz de Tenerife, Las Palmas de Gran Canarias, Baleares, Alicante, Madrid or Barcelona. *Vox* (right-wing) is the political party with the most radical approaches to immigration (deportation of illegal immigrants, deportation of legal immigrants with crimes, harsher penalties for immigration crimes ...).

Each of these facts can take fuzzy values: high, medium, low, indicating each of them the level of influence.

4.3. Influence of key elements.

Once the expert has identified the facts that have influence, he has also indicated the following conditions:

1) *PSOE* (left wing): You will get members of parliament in all territories. The *PSOE* voter will go to vote because in the last Autonomy Elections to the Parliament of Andalucía, 500,000 people did not go to vote and in the end governs the right wing. You will get more Members of Parliament as a consequence that more people will go to vote and from the loss of votes from *Unidas Podemos*.
2) The *PSOE* will win the 2019 General Elections. In Spain, the political party that wins the elections is not necessarily the political party that governs the country.
3) *Partido Popular* (right wing): You will not get or will get few Members of Parliament in the País Vasco and Catalonia. There will be a transfer of votes to *Ciudadanos* and *Vox*.
4) *Unidas Podemos* (left wing): Approximately, you will lose half of your current Members of Parliament. They were born as a regenerating political party to the traditional political parties (*PSOE* and *Partido Popular*) and increasingly resemble them. Some of its voters will go to the *PSOE*.
5) *Ciudadanos* (right wing): You will receive votes from the *Partido Popular*. It will have great success in Castilla y León.
6) *Vox* (righ wing): It is a new party that did not have parliamentary representation. It will obtain votes from the *Partido Popular* in territories such as Madrid, Murcia, Andalucía and Comunidad Valenciana due, in some cases, to immigration, and in others to the corruption of the *Partido Popular*.

Table 3. Degree of influence of factors in each territory.

Territory	Right wing Fragmentation	Catalonia's Independence	The defence of the rural world	Disenchantment of the *Partido Popular* voter	Immigration
Andalucía	Medium	Low	Low	Low	High
Aragón	Medium	High	Med	Low	Low
Asturias	Medium	Medium	Low	Low	Low
Cantabria	Low	Low	Low	Low	Low
Cataluña	Low	High	Low	Medium	Medium
Ceuta	Low	Low	Low	High	High
Castilla La Mancha	Medium	Low	Medium	Low	Low
Comunidad Valenciana	Medium	High	Low	Medium	Medium
Castilla y León	High	Low	High	Medium	Low
Extremadura	Medium	Low	Medium	Low	Low
Galicia	Low	Medium	Medium	Low	Low
Islas Baleares	Medium	Medium	Low	High	Medium
Islas Canarias	Medium	Low	Low	Medium	High
Comunidad del Madrid	High	Medium	Low	Medium	Medium
Melilla	Low	Low	Low	Medium	High
Murcia	Medium	Low	Low	Medium	High
Navarra	Medium	High	Low	Medium	Low
País Vasco	Low	High	Low	Low	Low
La Rioja	Medium	Low	Low	Medium	Low

In this phase, for each territory, the influence of each one of the factors that the expert has indicated in the final result will be determined, valuing the previous conditions (Table 3):

4.4. New Forecast: Historical data plus influencing factors.

In this phase, the analysis derived from the historical data of the dashboard is combined with the factors that the expert has indicated may have an influence on establishing a new forecast obtained from the number of Members of Parliament for each Autonomous Community, To simplify the reading of the article, it has been done at the level of Autonomous Community, territorial entity that groups provinces (Table 4):

In the next 2019 Spain General Elections, the *PSOE* (left-wing) will win, that the *Partido Popular* (right-wing) loses Members of Parliament to *Vox*, that the *Unidas Podemos* (left-wing) will lose up to half of its Members of Parliament that *Ciudadanos* (right-wing) will have little variation, and that *Vox* (right-wing) enters the Parliament with more than 20 Members of Parliament.

In the 2019 General Elections it is difficult for the CIS prediction to be fulfilled with respect to the left-wing (180 Member of Parliament) or for the right-wing to collapse to 135 Members of Parliament. It is forecast that the right-wing parties, *Partido Popular, Ciudadanos and Vox*, will get over 165 Members of Parliament, and the left-wing parties, *Partido Socialista and Unidas Podemos* will get over 155 Members of Parliament.

5. Analysis post electoral General Elections 2019 in Spain.

The results obtained in the General Elections of 2019 have been (Table 5):

Table 4. Forecast of the number of Members of Parliament by Autonomous Community.

Autonomous Community	Forecast Members of Parliament left wing	Forecast Members of Parliament right wing

Andalucía	28	33
Aragón	7	6
Asturias	4	4
Cantabria	3	2
Cataluña	16	11
Ceuta	0	1
Castilla La Mancha	12	9
Comunidad Valenciana	14	18
Castilla y León	15	17
Extremadura	5	5
Galicia	12	11
Islas Baleares	4	4
Islas Canarias	7	7
Comunidad del Madrid	13	23
Melilla	0	1
Murcia	3	7
Navarra	2	3
Pais Vasco	8	1
La Rioja	2	2
TOTAL	**155**	**165**

Table 5. Results of General Elections 2019 Spain.

	Member of Parliament	**Votes**	**%**
PSOE (left-wing)	123	7,480,755	26.68%
Partido Popular (right-wing)	66	4,356,023	16.7%
Ciudadanos (right-wing)	57	4,136,600	15.86%
Unidas Podemos (left-wing)	42	3,732,929	14.31%
Vox (right-wing)	24	2,677,173	10.26%

Participation was 75.75%, very high in reference to the previous Elections of 2016 and 2015 where it was 66.48% and 69.67% respectively.

As can be seen, practically the left and right blocks have obtained the same votes.

1) Sum of votes of the left-wing block: 11,213,684 votes.
2) Sum of votes of the right-wing block: 11,169,796 votes.
3) Difference: 43,888 votes in favour of the left block.

The comparison of the Members of Parliament obtained by each block in the General Elections of 2019, those predicted by CIS and those predicted by the model proposed by the article are (Table 6):

Table 6. Comparative results vs forecasting CIS vs forecasting proposed model.

	Result 2019	**Forecasting CIS**	**Difference**	**Forecasting model**	**Difference**
Left-wing	165	180	-15	155	+10
Right-wing	147	135	+12	165	-18

The CIS has failed to predict the Members of Parliament of the left block and has better approximated the Members of Parliament obtained by the right block. The model proposed in the article has better approximated the Members of Parliament obtained by the left block and has failed to approximate the Members of Parliament by the right block.

In the proposed model, a dashboard has been developed using the Qlikview electoral data tool since 2004. The article has shown that the analysis of data using Business Intelligence (BI) is not sufficient to make forecasts in 2019 since from the analysis of historical data in previous General Elections, the *Partido Popular* should have won in provinces such as Cantabria, Alicante, Madrid, Segovia, Valladolid, Zamora, Cuenca ..., on the other hand, the *PSOE* has won. The trend has been broken.

In the forecasting model proposed in the article, it has been forecasted correctly:

1) The *PSOE* will win the 2019 General Elections. In Spain, with a few exceptions, when the participation exceeds 70%, the PSOE wins. In the 2019 General Elections, the participation was 75.50% and the PSOE won the General Elections.
2) The *Partido Popular* will lose votes in favour of *Vox*. It is estimated that there has been a transfer of 1.6 million votes from the *Partido Popular* to *Vox*.
3) *Unidas Podemos* lose approximately half of the deputies to the *PSOE* and abstention. In the 2016 General Elections *Unidas Podemos* got 71 Members of Parliament and in 2019, 42 Members of Parliament (-29 Deputies). The *PSOE* obtained 85 Members of Parliament in 2016 and 123 Members of Parliament (+38 Deputies) in 2019. The *PSOE* has captured almost all the votes that the *Unidas Podemos* has lost.
4) The *Partido Popular* and Vox will not obtain Members of Parliament in the provinces of Alava, Vizcaya, Lleida, Tarragona and Girona (Pais Vasco and Catalonia).

5) The *PSOE* will obtain Members of Parliament in all provinces. Observing the results of the 2019 General Elections, it can be seen that the *PSOE* has effectively obtained Members of Parliament in all the provinces of Spain.
6) *Unidas Podemos* cannot get Members of Parliament in small provinces (from 1 to 5 Members of Parliament).

The proposed model has failed in:

1) *Ciudadanos*, more or less will maintain the Members of Parliament of the 2016 General Elections. In 2016 *Ciudadanos* had 32 Members of Parliament, in 2019 got 57 Members of Parliament. No forecasting or expert opinion, gave to *Ciudadanos* more than 50 Members of Parliament; it has been a surprise the 57 Members of Parliament it has obtained.

In the 2019 General Elections in Spain, the winners were the *PSOE* (which obtained 123 Members of Parliament when it previously had 85) and *Ciudadanos* (which went from 32 to 57 Members of Parliament). The losers were the *Partido Popular* (which went from 137 Members of Parliament to 66 in the last elections) and *Unidas Podemos*, which went from 71 Members of Parliament to 42.

With respect to the total number of Members of Parliament who have obtained the left and right wing, the model has failed mainly in the forecasting of the right wing block, where 165 Members of Parliament were predicted and has actually obtained 147 deputies. This has been due to the fact that there have been three political parties in this block dividing the vote. Never before has the Spanish elector, in a General Election, had so many options in right political parties with national implantation.

The fragmentation of the right-wing block's vote has mainly benefited the political party that came first in votes in the province, specifically the PSOE. The *Partido Popular* has lost almost three and a half million votes compared to the 2016 General Elections, where 1.6 million votes went to *Vox* and 1.4 million to *Ciudadanos*. The remaining were abstentionist voters. *Vox* took half a million votes from *Ciudadanos*. *Ciudadanos*, who obtained 3.1 million votes in the 2016 General Elections and one million more votes in the 2019 General Elections, have benefited from 1.3 million *Partido Popular* votes and high turnout.

Although the votes obtained by the two blocks hardly differ (0.5% difference), the left block has benefited greatly from the fragmentation of the vote of the right and the application of the D'Hondt rule: there have been certain municipalities where *Ciudadanos* and *Vox* have taken votes away from the *Partido Popular*, but they have not obtained enough votes to obtain at least one Member of Parliament. In total, the right-wing block has lost more than a million votes that have not been transformed into Members of Parliament. The *Partido Popular* has wasted 147,000 votes; *Ciudadanos* 164,000 votes; and *Vox* 689,000 votes. If the *Partido Popular*, *Ciudadanos* and *Vox* had participated together in the elections, they would have obtained an absolute majority with 177 Members of Parliament. In that scenario, the *PSOE* would have gone down to 108 Members of Parliament. And if *Vox* had not run, the Partido Popular would have obtained 98 Members of Parliament and 60 *Ciudadanos.*

Other issues highlighted in the article's model that have decisively influenced the 2019 General Elections are the following:

1) Catalonia's independence challenge. *Esquerra Republicana de Cataluña* (ERC) was the clear winner of the 2019 General Elections in Catalonia, rising from 9 to 15 Members of Parliament.
2) The defence of the rural world and emptied Spain. *Ciudadanos* has obtained 13 seats of the 66 Members of Paliament that were at stake in the small provinces (from 1 to 5 Members of Paliament). It has obtained representation in 19 of the 25 smallest provinces and in these provinces has obtained half of the 25 Deputies who have gone up from the General Elections of 2016 to 2019.
3) The disenchantment of the *Partido Popular* voter. The *Partido Popular* has lost almost 3.5 million votes in favour of *Ciudadanos* and *Vox.*

The main aspect by which the model has failed to forecast the Members of Parliament for the right wing has been not to assess in its entirety the fragmentation of the right wing, it was not thought that the *Partido Popular* would lose so many Members of Parliament in favour of *Ciudadanos.*

Although the CIS has obtained better data in its forecasting of the 2019 General Elections, in the electoral forecasts it has made before and after the April 28, 2019 General Elections, it has mistaken the forecast in other recent electoral processes. The elections that took place just before the 2019 General Elections were those of the Parliament of Andalusia in December 2018, the CIS indicated that the left wing clearly won, but finally the right wing won. The elections that have taken place after the General Elections of 2019, have been the Municipal, Autonomous and European Parliament Elections of 26 May 2019. An example will be the forecasting made by the CIS in the European Elections, where the CIS practically failed in almost all the forecasts of the Europarliamentarians that each political party

was going to obtain.

6. Conclusions.

Because in the field of data analysis there is usually no clear distinction between prediction and forecast, a mechanism has been proposed, in a preliminary way, to generate forecasts in complex problems, not predictable through estimation but conditioned by a series of specific events. This mechanism is based on forecasting based on the analysis of historical and current data plus the judgment of an expert who indicates which factors may influence the final result and to what extent. As future work we intend to improve the model using approximate reasoning and fuzzy logic mechanisms, in particular fuzzy deformable prototypes, to have a mechanism capable of adapting uncovered patterns from previous situations to a new and unknown one[2].

Acknowledgments

This work has been partially supported by FEDER and the State Research Agency (AEI) of the Spanish Ministry of Economy and Competition under grant MERINET: TIN2016-76843-C4-2-R (AEI/FEDER, UE).

Competing interests

The authors have declared that no competing interests exist.

References

[1] Stoetzer, L., Neunhoeffer, M., Gschwend, T., Munzert, S., Sternberg, S. (2019). *Forecasting Elections in Multiparty Systems: A Bayesian Approach Combining Polls and Fundamentals*. Political Analysis, 27(2), 255-262.

[2] Graefe A., Green K.C., Armstrong J.S. (2019). *Accuracy gains from conservative forecasting: Tests using variations of 19 econometric models to predict 154 elections in 10 countries*. PLoS ONE 14(1): e0209850.

[3] Huber, S. (2018). *The role of polls for election forecasting in German state elections*. Doctoral Thesis. Universität Passau, Deutschland.

[4] Graefe, A. (2018). *Predicting elections: Experts, polls, and fundamentals*. Judgment and Decision Making, 13(4), 334-344.

[2] The substantial variations that may have occurred between the data in this paper and the actual results are partly produced because this paper was sent on April 13, 2019 and the General Elections were on April 26, 2019. According to the polls a third of the voters decide their vote the last week before the day of the voting.

[5] Kaufman, A. R., Kraft, P., Sen, M. (2018). *Improving Supreme Court Forecasting Using Boosted Decision Trees*. Political Analysis, 1-7.

[6] De Rosa, R. (2018). *The use of big data in political-electoral communication. Forecasting the vote in the French presidential elections 2017*. Comunicazione politica 2/2018, 199-226

[7] Vaidehi, V., Monica, S., Mohamed Sheik Safeer, S., Deepika, M., & Sangeetha, S. (2008). *A prediction system based on fuzzy logic*. Proc. of the World Congress on Engineering and Computer Science, WCECS 2008.

[8] Atsalakis, George & K, Valavanis. (2013). *Surveying stock market forecasting techniques - Part I: Conventional methods. In Computation Optimization in Economics and Finance Research Compendium*. Nova Science Publishers.

[9] Atsalakis, G. S., & Valavanis, K. P. (2009). *Surveying stock market forecasting techniques–Part II: Soft computing methods*. Expert Systems with Applications, 36(3), 5932-5941.

[10] Qlik. "*2019 Magic Quadrant for Analytics and Business Intelligence Platforms*". Retrieved from https://www.qlik.com/es-es/gartner-magic-quadrant-business-intelligence, January 2019.

[11] *Rule D'Hondt*. Retrieved from https://es.wikipedia.org/wiki/Sistema_D%27Hondt. January 2019.

[12] Immigration statistics. Retrieved from http://www.poderjudicial.es/, December 2018

Citation: A. Lorenzo and J. Olivas. *A case study of Forecasting Elections Results: beyond prediction based on Business Intelligence.* Journal of Computer Science & Technology, vol. 19, no. 2, pp. 143-152, 2019.
DOI: 10.24215/16666038.19.e14
Received: April 15, 2019 **Accepted:** September 23, 2019.
Copyright: This article is distributed under the terms of the Creative Commons License CC-BY-NC.