International Journal of Production Management and Engineering

https://doi.org/10.4995/ijpme.2021.15300
Received: 2021-03-22 Accepted: 2021-05-04

PHYRON: Cognitive computing for the creation of an innovative Intelligence Experience Center

Ruiz, M.[a], Rodriguez, J.J.[b], Erlaiz, G.[c], Olibares, I.[d]

[a]ISEA S.COOP. (MONDRAGON Corporation), Spain.
[b]INDABA Consultores S.L., Spain.
[b]SARETEKNIKA Servicios Globales Postventa S. Coop., Spain.
[b]LANALDEN S.A., Spain.

[a]*mruiz@isea.eus*, [b]*jjrodriguez@indaba.es*, [c]*gerlaiz@sareteknika.com*, [d]*iolibares@sareteknika.es*

Abstract: This research presents the results of a project called "PHYRON: Cognitive Computing for the creation of an innovative Intelligence Experience Center", funded by the Basque Government (Economic Development, Sustainability and Environment Department). The project started in April 2019 and it will end in December 2021. Its main objective was to arrange an industrial research about cognitive computing. The main aim was the application of these systems for the development of an Intelligent Experience Center (IExC) to facilitate: i) enrichment of processes, products and services, in general client experiences, ii) automatic generation of technical predictions related to the product and the client behaviour through the exploitation of acquired knowledge, and iii) rationalization and automation of the processes that are involved in the after sale services both at technical and management level. The technological outcome presented in this paper is built using cognitive engines to enable learning from the client experience, and predictive models to anticipate client necessities.

Key words: cognitive computing, digitalization, digital transformation, predictive models, algorithms, big data.

1. Introduction

Digitalization is the central pillar of the current fourth industrial revolution that we are living now (Hagberg *et al.*, 2016). Digitalization is progressively penetrating in sectors such as industry, banking or retail. These sectors are related to "Industrial Internet or cyber-physical systems". In the past, added value in industry was based on what traditionally was understood as product and production, however, nowadays added value is related to data in the company. This fact has a direct impact on how competitiveness is understood (only thorough elements such as product design, quality or productive efficiency).

Moreover, the field based on data is feed by new topics connected to information and communication technologies, such as big data and cloud computing.

Due to digitalization industry and services are converging. This enables the entry of new agents with no productive capacities in the industry, and definitively the concept of manufacture industry is now being analysed.

2. Objective

In this paper the results of a Project called PHYRON are presented, in which the principal objective is

To cite this article: Ruiz, M., Rodriguez, J.J., Erlaiz, G., Olibares, I. (2021). PHYRON: Cognitive computing for the creation of an innovative Intelligence Experience Center. *International Journal of Production Management and Engineering, 9*(2), 103-112. https://doi.org/10.4995/ijpme.2021.15300

to arrange an industrial research about cognitive computing. The main aim is the application of these systems for the development of an Intelligent Experience Center (IExC).

The principal goal of this development is to facilitate technological improvements aligned with the business processes they support. In this case, technological advances for the improvement and efficacy of the call center service are: i) enrichment of processes, products and services, in general client experiences, ii) automatic generation of technical predictions related to the product and the client behaviour through the exploitation of acquired knowledge, and iii) rationalization and automation of the processes that are involved in the after sale services both at technical level and management level. The principal technological steps for the achievement of these goals are the definition of technological infrastructures for huge amount of data treatment, extraction, capture process and aggregation, and definition of predictive models. Technology explained in Section 5 (Results).

In this way, new knowledge about cognitive computing and its application in the industrial field is generated. An innovative product/service that combines infrastructure, software, and necessary tools for the development and implantation of an innovative Intelligent Experience Center (IExC) is the final result of the project.

PHYRON is based on the development of an IExC for after sale services and repairing of home appliances. The Consortium behind this Project is composed by four companies:

SARETEKNIKA Servicios Globales Postventa S. Coop., as an expert company in the after sale global services in the field of home appliances (Spain, Portugal, Andorra).

INDABA Consultores S.L., as an expert company in the design and implantation of Information Systems and Knowledge Management based on computing technologies.

LANALDEN S.A., experts in service through contact centre technology for process improvement, and technological innovation.

ISEA S.Coop., R&D institute of MONDRAGON Corporation. Agent of the Basque Network for Science, Technology and Innovation. They are experts in the development of technological research projects and launching of new activities.

SARETEKNIKA and LANALDEN modelled the solution for making it accessible for service and industrial companies. This new product/service is built with cognitive engines that used predictive models to anticipate customer needs. LANALDEN was the company in which the validation and pilot test was arranged. Finally, ISEA leads the work packages related to knowledge dissemination and new business model definition.

3. Literature Review: Artificial Intelligence (AI)

When the term AI appeared in the 1950s, there was access to a limited amount of data, and therefore AI did not evolve in the way and with the speed that was envisioned. Nowadays, however, there is a big amount of data which facilitates the development of AI (Forrester, 2017).

AI technologies need a large amount of information to execute algorithms (Engelmore, 1987), and consequently obtain a result which is close to the reality. These algorithms based their complex calculus on data. Consequently, data is turned into the key factor for the success of the process.

The objective of AI is not to replace the human being; it is to facilitate the analysis of the volume of data necessary for making intelligent decisions (Steels, 1993). In order to carry out these skills, a set of mathematical and statistical techniques is required to develop algorithms for the improvement of tasks based on experience and learning (machine learning). Therefore, the access to a greater volume of data facilitates the learning of AI through commands to improve decision making. AI is therefore the responsible for providing the necessary intelligence for the extraction of relevant knowledge and information (Bringsjord and Schimanski, 2003; Agrawal *et al.*, 2017).

As a previous step to any interpretation of data from AI perspective, it is essential to ensure the high quality of data (Bond and Gasser, 2014). For this purpose, a set of infrastructures and technologies is required to provide solutions and process huge data sets (structured, unstructured and semi-structured data). Data to be processed needs to be based on relational data rules (easily processable) in order to be considered structured data. Unstructured data concept, however, refers to data that has no value until it is ordered and classified (comments on social networks, images, audios, sensor data, etc).

3.1. Application of AI

In the following years disruptive business models will emerge and this will force companies to understand that digital transformation more than a trend is considered an essential point to remain competitive (PwC, 2017).

In the following list the advantages of AI are presented:

- *Resource and time optimization* through process automation and repetitive tasks.
- *Cost reduction* for the long period.
- Increase in *productivity and efficiency*.
- *Decision making* improvement due to efficacy and the increasing speed.
- Creation of *new activity lines* or business opportunities.
- *Improvement of client satisfaction* through the obtaining of different perspectives to predict their preferences and offer a better and personalized experience.
- Application of *human abilities through computational services*.

Machine learning (Zhang, 2020) and AI technologies are helping marketers to correlate and synthesize variables from different sources, identity patterns of behaviour, and inter consumer interest or purchase intent. This will result into the definition of individualized one-to-one marketing strategies that aim to achieve the highest level of personalization. Therefore, companies must define the best doing tasks. In this way, definition of responsibilities, roles and processes for efficiency maximization will be more efficient (Strong, 2016).

The collaboration between humans and machines is questioned by some experts; this relationship is understood as the transformation of tasks previously developed by humans into tasks developed by robots. Experts predict that up to 40% of jobs in the US and up to 30% in the UK can be lost due to AI. In addition, this trend will also lead to creation of new jobs in the fields related to soft skills. The reorganization of existing skills and the large scale hiring of new workers will be also a direct consequence of AI (PwC, 2017).

The new scenario resulted from AI technologies will move on from the traditional focus on low-value activities towards high value activities. The widespread automation of low-skilled jobs will change the business model of companies; they will be focused on workforce with higher productivity and fewer staff. During this transition, the main pillar will be the development of human capital and the application of new and more agile hiring strategies that reflect the change in the necessary skills paradigm for the competitiveness. Moreover, the automation of processes through the use of machines will enable the necessity of human skills, such as intuition, critical thinking, or creativity.

However, there are also some barriers related to the implementation of IA technologies:

- Fear of change
- Long periods of time for integration
- Adaptation to the use of advanced technology
- Lack of specialized talent in AI technologies
- Lack of knowledge and maturity in the field of solutions and providers
- Legal problems
- Not standardized impact

3.1.1. Machine Learning (ML), Deep Learning (DL) and Client Experience

In a more specialized level under Machine Learning (Zhang, 2020) is Deep Learning. It facilitates learning processes through computer models acting like neural networks in human brain. Its added value is the ability to learn from data through algorithms (Goldberg and Holland, 1988). In this way, rules of behaviour under algorithms replicate human rules of behaviour.

The application of deep learning (Rusk, 2016; LeCun *et al.*, 2015) enables the interpretation of the reality trough image recognition or natural language analysis. These functionalities are known as cognitive services. These services are focused on the replication of sensitive abilities of human beings, which makes them essential in the field of Customer Experience.

The application possibilities of this cognitive technology are very extensive (Hollnagel and Woods, 1983). The Association for the development of Customer Experience identified a 64% of implementation of customer voice programs (VoC) in companies. The objective of these programs is to facilitate the reception of client messages, identifying strengths and weaknesses of the service. The information about the client experience is extracted from the interactions in different channels, such as RRSS, emails, recordings of Contact Center calls, security videos, etc. Cognitive services would make possible the increase of the efficiency of these programs for the spread of analysed unstructured data and understanding through language recognition and image recognition algorithms.

As a result, it could be stated that nowadays AI market is a trending sector with support of worldwide technological experts. In this way, new customer demand is more efficiently responded. The companies leading this development offer a wide variety of AI solutions (chatbots, virtual assistants, machine learning algorithms…). They also offer the possibility to integrate other products such as CRM to increase the power of the solutions in a joint and linked strategy.

The anticipation of certain future scenarios can be very enriching for decision-making. Predictive models enable the anticipation of events for trust based relationship with clients, to personalize interactions with them and to anticipate to their needs. Predictive analytics are frequently used for both private and public sectors.

The analysis of customer lifecycle results in different solutions for direct applications. In the case of sale process, additional product recommendation models (cross-selling) or predictive demand models are widely used solutions among International and Spanish companies. The differential value of this technique is the use of segmentation models. They facilitate the perception of the existing logic behind two different clients with similar behaviours.

3.1.2. Intelligence Experience Center

During the next years, consumers will regularly interact with cognitive computing-based services. Cognitive systems and artificial intelligence in general, facilitates the management of huge volumes of data in companies. Cognitive systems are changing the way in which people and systems interact (High, 2012). This has a visible impact in the communication process settled down in contact centres.

Converge and technological advance, such as Big Data, Cloud Computing or the Internet of Things (IoT) are the levers for the change in this context, and they are integrated in the system to respond to the new customer needs. They are driving growth and innovation in AI. These technologies will enable management of both information and strategic decision making through AI, to elevate customer service to a more personalized level (PwC, 2018).

The new emerging customer is a more demanding and informed consumer who places new demands on companies and expects a greater degree of customization from companies. Different studies prove that the biggest challenge for Spanish companies is the personalization of products and services to respond to client real needs.

4. Methods

4.1. General Process

Big Data projects generally require 2 processes: i) design of the solution, and ii) implementation of the solution (Wu *et al.*, 2013; Mousannif *et al.*, 2014; Katal *et al.*, 2013; Jin *et al.*, 2015). These processes are divided in 5 phases:

PHASE 01 Data Sources. This first phase requires the identification of data sources, type of source (relationship database, spreadsheets), file format (.txt, .doc, .xls…), or data character (data structure). In a Big Data environment, it will be important to determine all the data sources that are required for the execution of the project.

PHASE 02 Data Ingestion. After the identification of data sources and their presentation, the next step is to select the techniques for extracting data, it can be extracted from a traditional ETL (Extract, Transform and Load or "extract, transform and load" process). This technique could be applied on traditional structured data, Crawling techniques to extract data from web pages, design interoperability services, application of different APIs or Application Programming Interface, subroutines, functions or specific procedures.

PHASE 03 Data Storing. In this phase, the required data storage solution is designed and implemented (relational database, datawarehouse or a non-SQL database, sized to the project requirements, in terms of volume, etc.).

PHASE 04 Data Processing. In this phase, the data processing or intelligence application techniques required in the project will be designed and implemented. Initially, data mining techniques will be applied. Data processing is arranged using tools such as Weka, and observing the results obtained. After the study, the algorithms that offer the best results can be implemented. The need to apply Natural Language Processing (NLP, Natural Language Processing) and semantic web techniques was also analyzed.

PHASE 05 Data Analysis. Finally, in this phase, graphical and usable interfaces were developed to show the outcome of the project to the users.

4.2. Methodology

Considering this general and standard methodology, PHYRON is developed during 33 months (from 2019 to 2021), and the main tasks of the project are three:

Table 1. Principal tasks for the execution of the methodlogy (own source).

Task	Name	Duration
01	Technological Research, case of uses definition, and project requirements	2019
02	Predictive models design and development	2020
03	Solution implantation, validation and result exploitation	2021

In the following Figure 1 the detail of the methodology is presented and in the next paragraphs each phase will be explained:

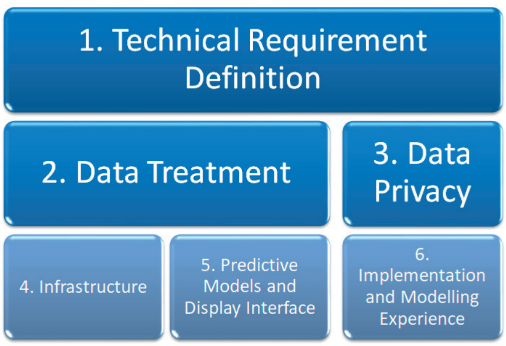

Figure 1. Principal phases.

1. Technical Requirement Definition:
1.1 Generation of new knowledge in Computer Science
1.2 Definition of the cases of use
1.3 Collection of security and privacy requirements
1.4 Definition and collection of security and privacy requirements
1.5 Collection and definition of the requirements for the design of the infrastructure
1.6 Collection and definition of requirements for the predictive models

This task was focused on the generation of new knowledge in Computer Science. The main objective was to have a general perspective about machine learning techniques and their application to industrial environments. In addition, the selection and description of the cases of use was done. This is the pillar for the requirements of data sources, security, infrastructure, and predictive models. Security refers to the anonymization of data. In order to define the infrastructure is necessary to describe hardware and software sets. Extraction, storage and process requirements were defined. Regarding predictive models, the most important objective was to facilitate decision making through the identification of entry parameters. Moreover, the requirements for graphical and attractive interfaces were defined.

2. Data Treatment:
2.1 Standardization and homogenization
2.2 Data Sources exploitation and analysis
2.3 Characteristics extraction
2.4 Design of the process for extraction, storage and aggregation
2.5 Implementation of the process
2.6 Validation of the process

The first step is related to the quality control applied to data, data source standardization was compulsory. Data source analysis was done based on the definition of the cases of use. In this way, available variables were selected for their use in the predictive models. This analysis showed that more variables than expected were needed. Cleaning and processing treatments were applied to variables to facilitate their use. Finally, the process for extraction, capture and processing was developed through the use of the engine for anonymization. For the validation, selected variables were exported step by step. Once the validation was done all the variables were exported.

3. Data Privacy:
3.1 Design of the process for data anonymization
3.2 Implementation of anonymization process
3.3 Implementation of infrastructure security

Data managed in this process belongs to particular users (individuals). Consequently, it was necessary to implement the anonymization process to guarantee its safe use in different units of storage. Anonymization engine made the variables invisible acting on its identity attribute. Security of the infrastructure was implemented in parallel with validation.

4. Infrastructure development:
4.1 Installation of the software
4.2 Validation of the infrastructure

Installation was dependant on the requirements defined in the first stage, and it was composed of different software sets to satisfy different functions such as extraction and data capture, storage, resource management, and predictive model generation through machine learning algorithms. Apart from the installation, configuration of each of the software sets facilitated validation of the infrastructure.

5. Predictive Models and Display Interface:
5.1 Predictive Models Design
5.2 Design of the display interface
5.3 Development and implementation of predictive models
5.4 Implementation of the display interface
5.5 Adaptation of predictive models for their execution in real time
5.6 Validation of the predictive models and display interface

Firstly, existing correlations between data were identified to simplify the complexity of the data set, as well as to identify how these variables could be enriched. When data was analysed, inputs and outputs for the final result were classified. The final result was defined according to the requirements in the first stage.

In that sense, when defining predictive models it was essential on the one hand the scalability of the variables, and on the other hand, model parallelization. These characteristics enabled the future integration of new parameters and future adaptation of variables to new necessities.

For the display interface design *User Centered Design* philosophy was followed and *mockups* were used. Moreover, information architecture and interaction were executed to optimize user experience.

Predictive models were validated independently. Secondly, display interfaces were validated. Finally, usability proofs were arranged to identify the possible mistakes related to design and development.

6. Implementation of IexC PHYRON and experience modelling:
6.1 Handbook and protocols for taking action
6.2 Pilot Test
6.3 Identification of failures

This phase was based on the definition of contents to facilitate the use of developed tools. A training program was defined for workers from LANALDEN. A real case study was arranged to validate the tool, for a high quality failure identification phase.

5. Results

The main objective of this Project was to analyse new profiles and necessities of final users of SARETEKNIKA. The basis to extract this information was the available information in SARETEKNIKA about usual casuistry and actors involved in their after-sale service processes.

This process was necessary to guarantee the total functionality of the system to solve real cases in after sale service. The main results could be defined as following: i) specifications of the architecture and extraction system, ii) anonymization process, and iii) infrastructure construction, and predictive models.

a. Specifications Architecture and Extraction System

During the first year, the extraction and processing system was defined and implemented. A selection was initially made for all the information collected, and considering relevant data aligned with the established treatment objectives.

The quality of the hidden knowledge to be discovered depended on both Machine Learning algorithm or the technique used and the quality of data to be treated. It is necessary to have a coherent raw material reception process to obtain consistent conclusions.

Then, transformation and cleaning of data for the modelling tools was developed. Techniques to ensure the quality of data were used:

- *Cleaning*: This phase was focused on correcting data anomalies that can result in inaccurate conclusions. These anomalies can result from missing data or from values that do not fulfil expected behaviour. It is important to identify the origin of these problems and their distribution before applying different corrective measures (statistical techniques to solve them).
- *Transformation*: When the predictive power of the attributes is not very clear, it is necessary to build new attributes by applying some transformation to the original attributes (grouping, separate dates to integers, converting categorical values into numbers).
- *Dimensionality reduction*: Select the appropriate set of attributes for the specific task to be performed. It is carried out based on statistical techniques combined with empirical methods. It will facilitate optimization of analytical models.

In the following Figure 2 the principal entities related to the notifications of failure could be seen:

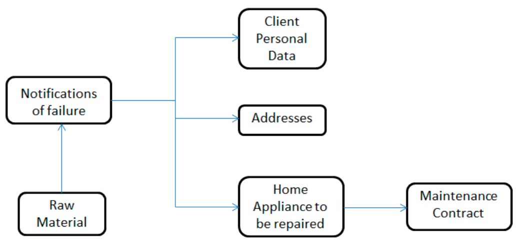

Figure 2. Principal entities related to the notifications of failure (own source).

b. Anonymisation Process

Initially, existing identifiers in Automatic Identification Systems (AIS) were eliminated from the PHYRON data source system. It was the first step to avoid individual identification through PHYRON data treatment.

- Due to the objective pursued in PHYRON, most of the sensitive attributes about clients and existing in the original data source are not necessary: Bank, IBAN, Bank card, contract amounts and repairs.
- Only two attributes have been maintained: "CHARGED", and "PAYMENT_DATE".
- Regarding the quasi-identifiers that have been used (Postal code, Town, Province), K-Anomymization techniques have been applied to achieve an acceptable level of anonymization considering affected users.

c. Infrastructure Development

PHYRON infrastructure is composed of four main modules: i) technological infrastructures for huge amount of data treatment, ii) extraction, capture, process and aggregation, iii) advanced analytics and predictive models, and iv) methodology and protocol for taking action.

COMPONENT 1: Technological Infrastructures for huge amount of data treatment

The technological infrastructure functioned as the basis to collect and process data to create predictive models for the personalized prediction of each client situation.

The adequate dimensioning of the technological infrastructure is a key factor to arrange the necessary data processing. An oversizing would increase the associated costs due to both the material cost and its management. The oversizing has no improvement in the processing time for predictions. Failure to meet storage and processing demand will decelerate the project development.

The software needed to manage and process data and generate models to be run in real time. It is a key factor for the sizing of the technological infrastructure.

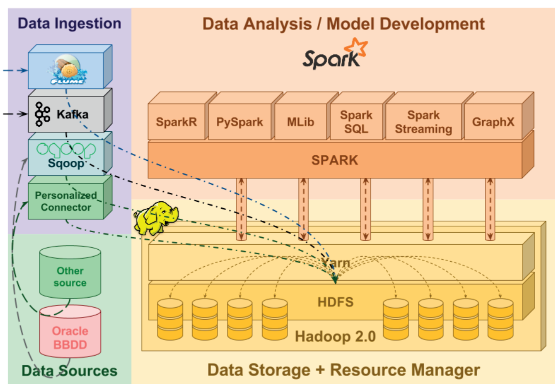

Figure 3. Software and data flow in PHYRON (own source).

PHYRON technological infrastructure is composed by software sets which belong to four different families: i) data sources, ii) data feed, iii) data collection and resource management, and iv) data analysis and model generation.

The implemented technological infrastructure will collect data for the predictive models (without prior treatment). Data flow shown in the following Figure 3 will use Spark Streaming to process data in real time.

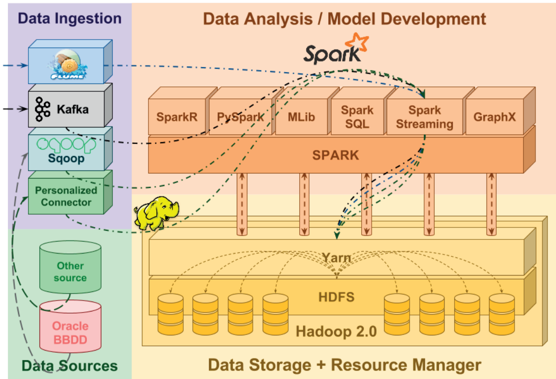

Figure 4. Software and data flow for streaming process (own source).

One of the fundamental aspects to implement a technological infrastructure is its ability to scale. The storage and processing needs will change depending on both the case of use and the amount of data generated. The original design and implementation is done for previously treated data set and processing method.

In addition, depending on possible future needs, it is desirable that the software offers integration with other sets of software packages for additional functions. In this sense, open source software offers the necessary independence for flexibility when integrating other software packages or developing connectors to facilitate integration.

COMPONENT 2: Extraction, capture, process and aggregation

PHYRON is based on data extraction technologies, both structured and unstructured data. The main challenge is this sense was to investigate an adjusted solution for each set of structured data. This concept of “data” refers to available and previously stored data in relational storage systems. Data generated in real time.

Moreover, different unstructured data is necessary as input for the predictive models. The solutions proposed for each type of unstructured data will contemplate the origin, frequency and, possible anomalies in the system. This results in the necessity to use different techniques depending on the data.

A clear example of the pre-processing of unstructured data is the data collected through free text fields specified by the technicians. These fields facilitate the freely definition of the causes for failure or service provided, resulting in a text that is easily understood by human being, but unintelligible for information systems. It could not be used without prior pre-processing and extraction of characteristics.

Apart from the ML algorithms selected, the quality of the hidden knowledge is also influenced by the quality of the data treated. When inconsistences are found in the flow, the conclusions extracted from the system are inconsistent.

COMPONENT 3: Advanced analytics and predictive models

PHYRON aimed to develop different techniques that enable the development of predictive models that can be used “online”, although their generation will be “offline”. The novelty arranges from the holistic approach of integrating the design and development of the techniques. Finally, developed tools were validated with real use cases.

During the last year of the project, the design of the predictive models will be developed, in parallel to the visualization interfaces creation. For the design phase it is necessary to contemplate both the result of the project, and the way to present the result.

COMPONENT 4: Methodology and Protocol For Taking Action

The fourth component of the IExC PHYRON was based on the elaboration of methodologies and documentation to collect the principal guidelines to facilitate an efficient customer experience, and optimizing the use of PHYRON.

In the same way, protocols for taking action in a personalized way and services associated to each of the contact phases with the client were defined. Training plans for the last year were designed.

d. Predictive Models Development

This process was developed in parallel with the design of the display of the results. It is necessary an alignment between "what to obtain" and "how to present it". In this way, added value in the process is increased.

This process is developed in an iterative form, an initial design is implemented and its outcomes are used for the redesign and optimization of the solution.

During this phase initial requirements were defined and it was especially relevant to follow them to guarantee an optimum integration degree (Steyerberg *et al.*, 2001; Kuhn, 2008; Biecek, 2018).

Variables that could be considered predictors were identified in a personalised way for each of the cases of planned use. This analysis was based on: i) familiarization with data, ii) identification of data quality, iii) identification of initial ideas, and iv) detection of data sets to form hypotheses about hidden information. After this process, the final result was the dataset to begin with the predictive model.

Once the infrastructure was assembled and data treatment process was designed and implemented, the predictive models were designed both for their execution based on historical data and their execution in real time.

Firstly design and development of models oriented to the existing possibility of repairing and the identification of the necessary materials for the repairing were done.

Conclusions from the subsequent model and its tests may lead to the convenience of reformulating its design, modifying the pre-processing of certain variables, or even incorporating new variables discarded in a first moment.

Results obtained from the models bring an improvement in the classification with respect to the distribution of variables.

6. Conclusions

The final result of this Project is a product and a new service related to cognitive computing: i) cognitive platform based on the application of machine learning algorithms for the development of predictive models for consumer behaviour analysis, and ii) new service based on cognitive services related to Intelligent Experience Centre.

Regarding process optimization it is necessary to mention that time and resources will be extremely improved due to PHYRON solution. Costs in the long period, productivity and operative efficiency, decision making processes, new activity lines, client satisfaction, and application of human abilities through automated computational services will be improved.

The result of the Project is the design of a cognitive system for an Intelligent Experience Center. This platform will be offered to the service sector using a commercial brand to convey this technological solution.

Acknowledgements

We would like to thank the Basque Government for their support in the development of this project. Special thanks to the Economic Development, Sustainability and Environment Department.

References

Agrawal, A., Gans, J., Goldfarb, A. (2017). *What to expect from artificial intelligence*. MIT Sloan Management Review.

Biecek, P. (2018). DALEX: explainers for complex predictive models in R. *The Journal of Machine Learning Research*, *19*(1), 3245-3249.

Bond, A.H., Gasser, L. (Eds.). (2014). *Readings in distributed artificial intelligence*. Morgan Kaufmann.

Bringsjord, S., Schimanski, B. (2003). What is artificial intelligence? Psychometric AI as an answer. In *IJCAI*, pp. 887-893.

Engelmore, R.S. (1987). Artificial intelligence and knowledge based systems: origins, methods and opportunities for NDE. In *Review of Progress in Quantitative Nondestructive Evaluation*, pp. 1-20. Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-1893-4_1

Goldberg, D.E., Holland, J.H. (1988). Genetic algorithms and machine learning. *Machine Learning 3*, 95-99. https://doi.org/10.1023/A:1022602019183

Hagberg, J., Sundstrom, M., Egels-Zandén, N. (2016). The digitalization of retailing: an exploratory framework. *International Journal of Retail & Distribution Management*. *44*(7), 694-712. https://doi.org/10.1108/IJRDM-09-2015-0140

High, R. (2012). The era of cognitive systems: An inside look at IBM Watson and how it works. *IBM Corporation, Redbooks*, 1-16.

Hollnagel, E., Woods, D.D. (1983). Cognitive systems engineering: New wine in new bottles. *International journal of man-machine studies*, *18*(6), 583-600. https://doi.org/10.1016/S0020-7373(83)80034-0

Jin, X., Wah, B.W., Cheng, X., Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, *2*(2), 59-64. https://doi.org/10.1016/j.bdr.2015.01.006

Katal, A., Wazid, M., Goudar, R.H. (2013). Big data: issues, challenges, tools and good practices. In *2013 Sixth international conference on contemporary computing (IC3)*, 404-409. IEEE. https://doi.org/10.1109/IC3.2013.6612229

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1-26. https://doi.org/10.18637/jss.v028.i05

Forrester (2017). "Predictions 2018: The Honeymoon For AI is Over".

LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Mousannif, H., Sabah, H., Douiji, Y., Sayad, Y.O. (2014). From big data to big projects: a step-by-step roadmap. In *2014 International Conference on Future Internet of Things and Cloud*, 373-378. IEEE. https://doi.org/10.1109/FiCloud.2014.66

PwC. (2017). UK Economic Outlook. Recuperado de http:// www.pwc.co.uk/economics

PwC. (2018). "Bots, Machine Learning, Servicios Cognitivos. Realidad y perspectivas de la Inteligencia Artificial en España, 2018".

Rusk, N. (2016). Deep learning. *Nature Methods*, *13*(35), 35-35. https://doi.org/10.1038/nmeth.3707

Steels, L. (1993). The artificial life roots of artificial intelligence. *Artificial life*, *1*(1,2), 75-110. https://doi.org/10.1162/artl.1993.1.1_2.75

Steyerberg, E.W., Harrell Jr, F.E., Borsboom, G.J., Eijkemans, M.J.C., Vergouwe, Y., Habbema, J.D.F. (2001). Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of clinical epidemiology*, *54*(8), 774-781. https://doi.org/10.1016/S0895-4356(01)00341-9

Strong, A.I. (2016). Applications of artificial intelligence & associated technologies. *Science [ETEBMS-2016]*, *5*(6).

Wu, X., Zhu, X., Wu, G.Q., Ding, W. (2013). Data mining with big data. *IEEE transactions on knowledge and data engineering*, *26*(1), 97-107. https://doi.org/10.1109/TKDE.2013.109

Zhang, X.D. (2020). Machine learning. In *A Matrix Algebra Approach to Artificial Intelligence*, 223-440. Springer, Singapore. https://doi.org/10.1007/978-981-15-2770-8_6

