ISSN:0975-9646

Alessandro Massaro et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 9 (1) , 2018, 4-9

Web App for Dynamic Pricing Modeling in Automotive Applications and Data Mining Analytics

Alessandro Massaro[#1], Angelo Galiano[#2],Giacomo Fanelli[#3], Bachir Bousshael[#4], Valeria Vitti[#5]

[#]*Dyrecta Lab, Research Institute*
Via V. Simplicio 45, 70014 Conversano (BA) - Italy

Abstract— **In this paper the authors propose a web application oriented on business to business (B2B) and internet automotive marketing. The study case is related to an industrial project concerning tires recognition application and real time comparison of prices of competitors. By means of a smartphone camera able to identify and classify car plates, the system provides different associations of compatible tires for a specific car model and the best prices. The frontend system is based on Optical Character Recognition (OCR) method and the backend works on automatic extraction of competitor prices from different databases and from internet network. The dynamic pricing is due to the best price choice related to period of car plate recognition. The price choice is supported by data mining processing performed by a k-Means workflow. The paper summarizes the results of a research industrial project. The used approaches could be used for different automotive applications involving market of different car parts.**

Keywords— **Web Application, Mobile Marketing, OCR, Tires detection, Dynamic Pricing, Data Mining, K-Means.**

I. INTRODUCTION

Different studies focused the attention on dynamic pricing in marketing and business [1]-[6]. The strategy to adopt changes with the application and with the typology of market. E-commerce and in general internet are powerful tools useful to adopt a strategy [7], to decrease price competition [8] and to provide customer motivation [9]. For this purpose internet channels could be data sources useful for real time monitoring of prices. The goal of the study is the matching between software tools oriented on online dynamic pricing and mobile ones able to acquire information about products to analyze. In fact the mobile technologies represents a "smart way" to recognize objects stored in a database. In particular, by considering automotive applications, some technique [10]-[12] could recognize care plates in order to attain data about car models, car proprietary, tires compatibility and so on. An approach used to recognize car plates is the Optical Character Recognition (OCR) which could be implemented by using techniques such as Automatic Number Plates Recognition (ANPR) [13] and suitable libraries as OpenCV [14]. According to these topics has been developed the conceptual model of Fig. 1, which is summarized in the following main requirements:

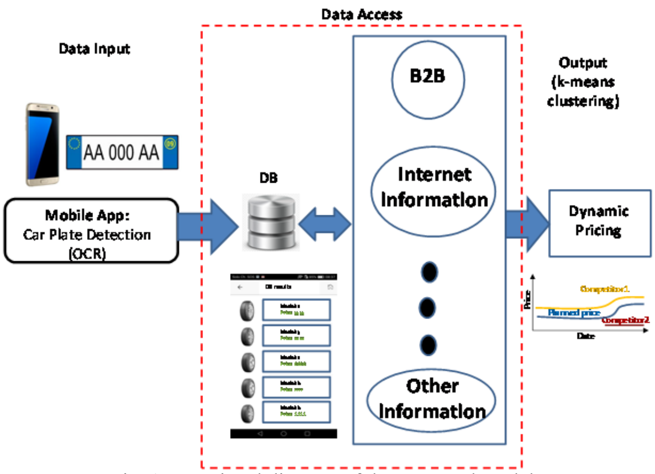

Fig. 1. Functional diagram of the proposed model.

i. An user detect the a car plates by means of a mobile app and by OCR algorithm (automatic OCR detection or manual data entry);
ii. The plate is recognized and a database provides a list of possible tires correlated with the car model;
iii. Different data available in internet or in B2B platforms about tires prices are compared: this comparison will provide a full dashboards of competitor prices;
iv. Data are collected and a data mining algorithm will provide a Decision Support System (DSS) about the best cluster and price to adopt (dynamic pricing updating with actual tires price and comparisons with competitors prices).

As shown in Fig. 1, the prototype system is structured in the following three main parts:

1) A frontend interface (mobile APP);
2) A backend system for data access and for the check of tires availability;
3) A data mining engine for dynamic pricing outputs.

Concerning data mining algorithms K-means approach is useful for data clustering thus providing a simple way to perform a DSS [15] by implementing data processing workflows by means of graphical user interfaces (GUIs) [16].

We will discuss in the next sessions the architecture and system design, the backend/front -end specifications of the prototype system, and, finally, a simulation of the dynamic pricing output.

II. SYSTEM DESIGN AND BACK/FRONT END FRAMEWORKS

In this section we will provide the whole design of the proposed project sketched in the basic architecture of Fig.1. In order to propose a full design scheme, different layouts of diagrams were illustrated. Unified Modeling Language (UML) has been used for the design layouts (UMLet tool). The design is divided in the following five layouts:

1. Use Case Diagram (UCD): in Fig. 2 is shown the UCD diagrams identifying the actors of the proposed model (customer is the actor having the car plate to detect, besides "visualization" is the user which utilize the OCR mobile application and the dynamic pricing analysis);
2. Activity Diagram (AD): in Fig. 3 is illustrated the flow chart concerning the main steps of the process of the tires detection, car model compatibility, and check of warehouse availability (main functions of the proposed model);
3. Class diagrams (CD): in Fig. 4 is sketched the structure of the prototype system by showing the system's classes, their attributes, operations and methods, and the relationships among objects (the CD include the classes of the backend and of the frontend systems);
4. Sequence Diagram (SD) B2B data access: Fig. 5 indicates the time sequence of the DB connection and query management process.
5. Sequence diagram (SD) input system: Fig. 6 explains the time sequence related to the mobile application (photo acquisition, manual data entry of the plates, OCR usability).

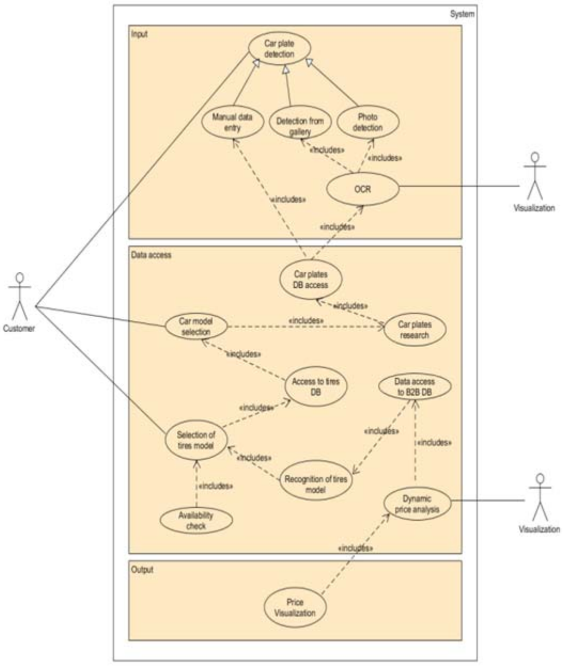

Fig. 2. Use Case Diagram of the project (UCD).

The full logic of the prototype architecture of Fig. 1 is developed in the diagrams of Fig. 2, Fig. 3, Fig. 4, Fig. 5 and Fig. 6. All diagrams were optimized after the testing phase therefore satisfy all the requirements.

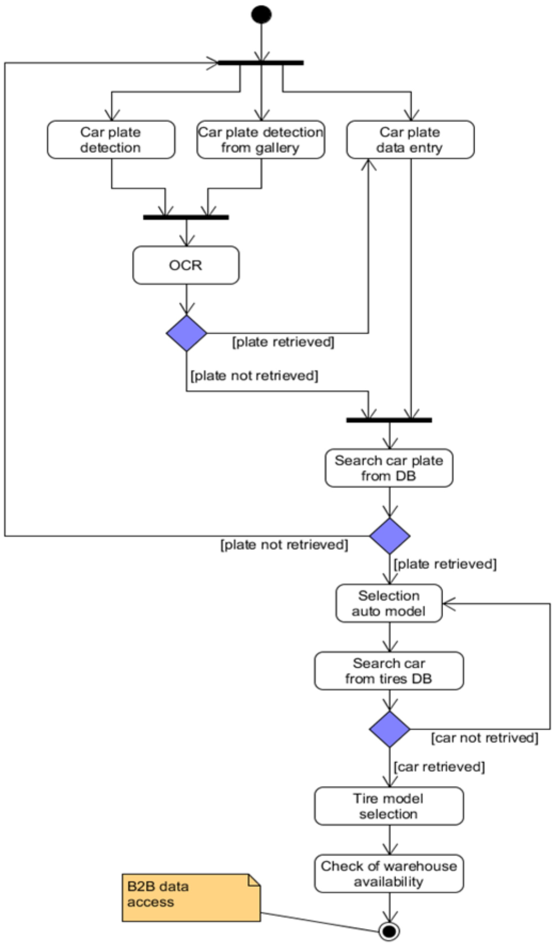

Fig. 3. Activity Diagram (AD).

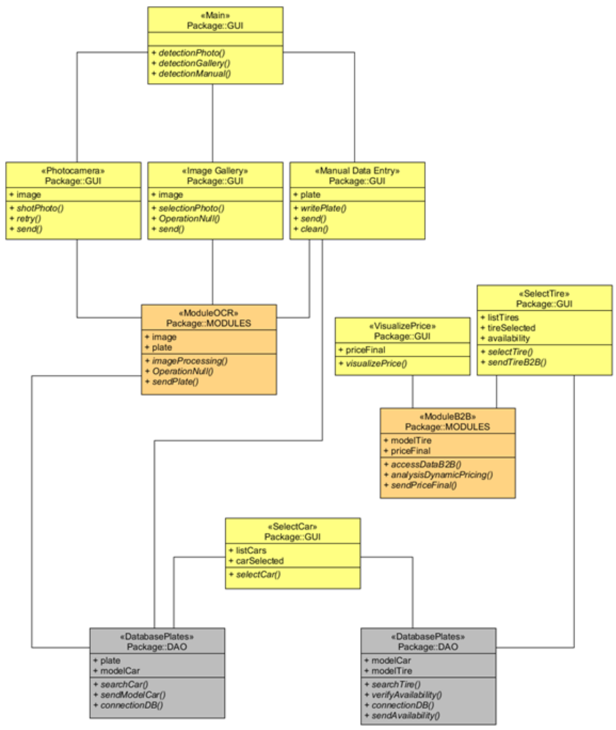

Fig. 4. Class Diagram (CD).

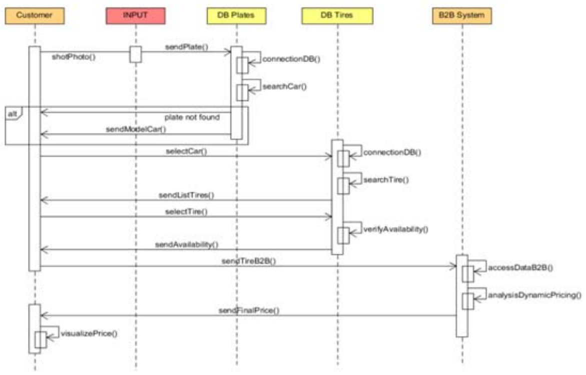

Fig. 5. Sequence Diagram: B2B data access.

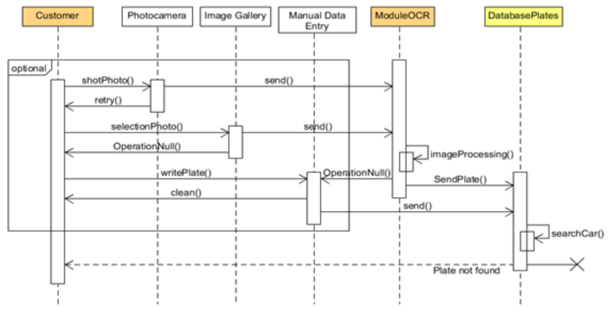

Fig. 6. Sequence Diagram: input system.

The framework used for the development of Android/Windows/IoS mobile APP is the platform integrating Ionic2, Angular 2 and Apache Cordova tools. The backend has been implemented by following the client/web service logic (http request/http response) accessing to the database by API. For the backend implementation it has been used Magento (based to Zend php framework) API and extended with new functionalities (for mobile app synchronization and price comparison). For the OCR plate recognition it has been developed an algorithm based on OpenALPR libraries. OpenALPR is an open source Automatic License Plate Recognition library written in C++ with bindings in C#, Java, Node.js, Go, and Python. The library analyses images and video streams to identify license plates. The output is the text representation of any license plate characters. OpenALPR requires the following additional libraries: Tesseract OCR and OpenCV. For the visualization of tires price of competitors has been implemented Rapid Miner tools reading an excel file, besides for data mining processing has been implemented an object workflow of KNIME Studio tools (clustering processing by k-Means algorithm). KNIME and Rapid Miner are open source data analytics, reporting and integrating various components for machine learning and data mining through their modular data pipelining concept (workflow). Graphical user interfaces (GUIs) allow assembly of nodes for data preprocessing (Extraction, Transformation, Loading –ETL-), for modeling and data analysis and visualization (reporting). Both the approach are high level programming tools and can be considered as SAS (Software as A Service) alternative. We observe that Rapid Miner offers a best Scattering Multiple reporting operator useful to compare together all the tires prices of more competitors.

III. Testing and Results

Different tests have been performed for the mobile APP. We describe below five main performed test:

- First test (Fig. 7): access to the system (login), photo menu (shot a photo, photo menu);
- Second test (Fig. 8): shot a car plate photo, processing (OCR algorithm), detection accuracy (probabilistic percentage of correct plates detection: in the figure is indicated an accuracy of the 90 % for the test plate);
- Third test (Fig. 9): manual data entry of the car plates, search by car function, search car by model function;
- Fourth test (Fig. 10): show tire details, show list of tires prices, advanced research function (by tires parameters provided by the database system);
- Fifth test (Fig. 11): buy/order function, tire model/ car model compatibility, selection of season for the tires choice.

The backend test has been focused on modifying of price field, on the data entry of competitor price and on the bidirectional linking with the frontend system. In Fig. 12 we illustrate a screenshot related to the adding procedure of a single price. The backend system is able to export data file in csv/excel format. These files will be processed by the data mining workflows. In order to check quickly the data mining workflow output, we have constructed the test Excel input file of Fig. 13: the file indicates randomic values of 44 types of tires model, the related prices of 4 main competitors, and the quantity sold by each competitor. In order to view the global price distribution of all tire models we used the “Scatter Multiple” function of Rapid Miner Tool (see Fig. 14): specifically we imported the testing file by means of the “Read Excel” object and by executing a basic input/output workflow (see inset of Fig. 14 where the object is linked by a “pipeline” to the input and to the output port). By running the basic workflow, and by selecting the “Scattering Multiple” function after the simulation, it is possible to plot all the tires prices (see graphical report of Fig. 14). By selecting the sold quantities will be possible to plot also the sold quantities distribution. The scattering multiple function is a first level analysis able to suggest a price according with competitoe trends. For a deeply analysis (second level analysis) we need of an advanced data mining algorithm. In this case of study we consider the k - means algorithm by setting k =5 (five cluster to evaluate). The algorithm was implemented by a KNIME object. This object outputs the cluster centers for a predefined number of clusters (no dynamic number of clusters, the dynamic price is to think as an updating of competitors dataset to analyze). K-Means performs a crisp clustering that assigns a data vector to exactly one cluster. The algorithm terminates when the cluster assignments do not change anymore. The clustering algorithm uses the Euclidean distance on the selected attributes. The data is not normalized by the “k -Means” object (if required, should be considered to use the "Normalizer" as a preprocessing step). For four competitors to analyze, five cluster are enough for data processing (five price levels matched with five level of quantities of sold tires). By decreasing the number of clusters the analysis could be poor for valid supporting decisions.

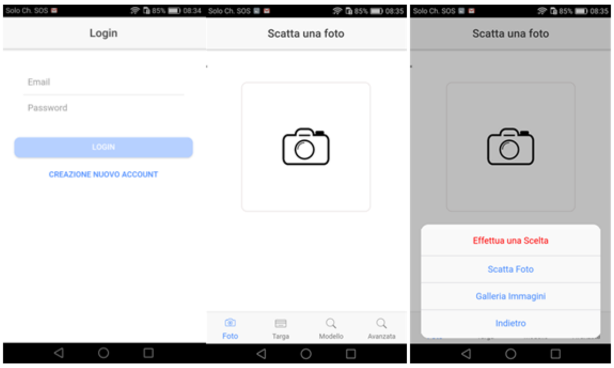

Fig. 7. First test of mobile APP.

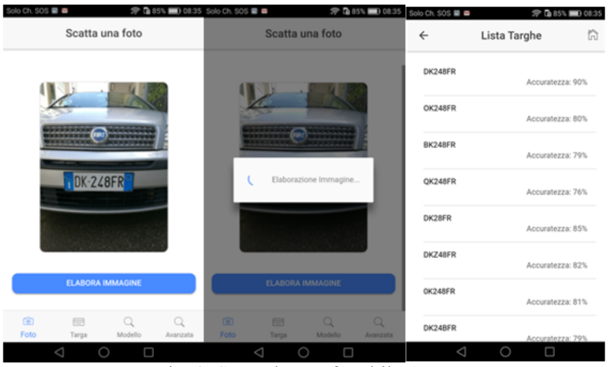

Fig. 8. Second test of mobile APP.

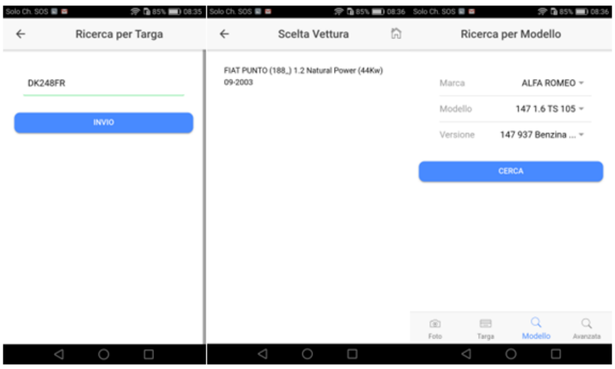

Fig. 9. Third test of mobile APP.

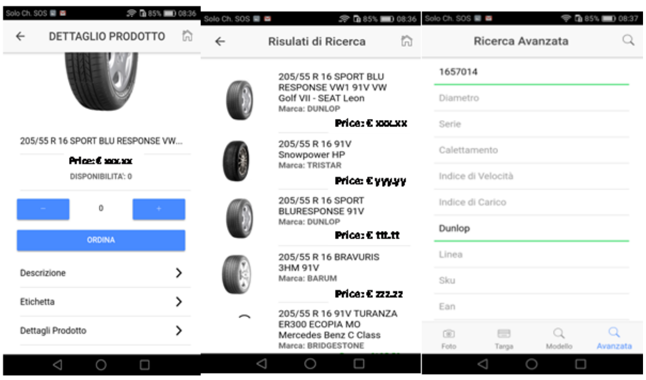

Fig. 10. Fourth test of mobile APP.'

In Fig. 15 is illustrated the KNIME workflow used for the k-Means analysis. The workflow is made by the following three objects: i) "Excel reader (XLS)" able to import the input file of Fig. 13; ii) "k-Means" object implementing k – Means algorithm; iii) "Scatter Plot" object which allows to plot clustering distribution.

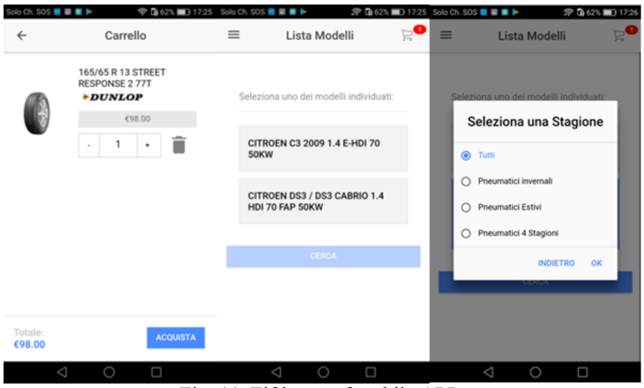

Fig. 11. Fifth test of mobile APP.

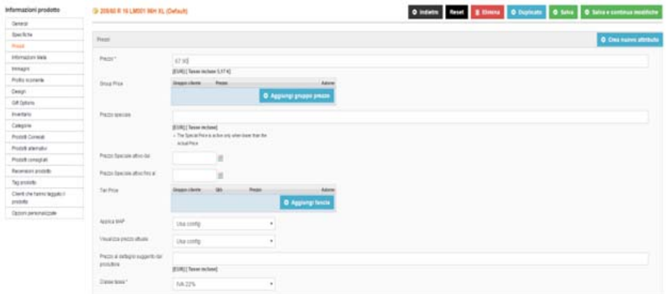

Fig. 12. Back end Test (add function of a tire price).

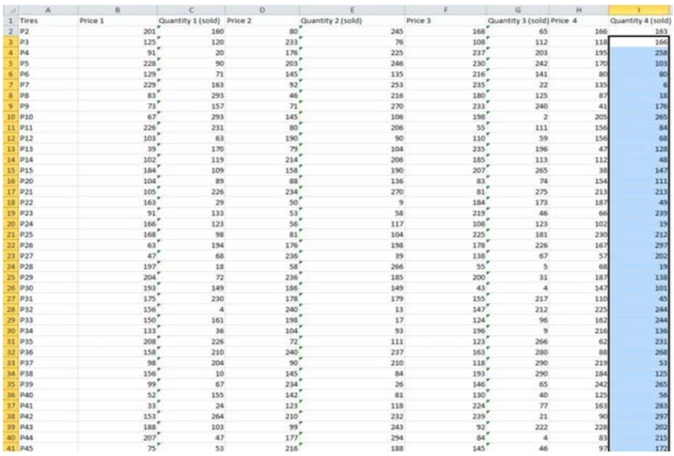

Fig. 13. Excel testing file.

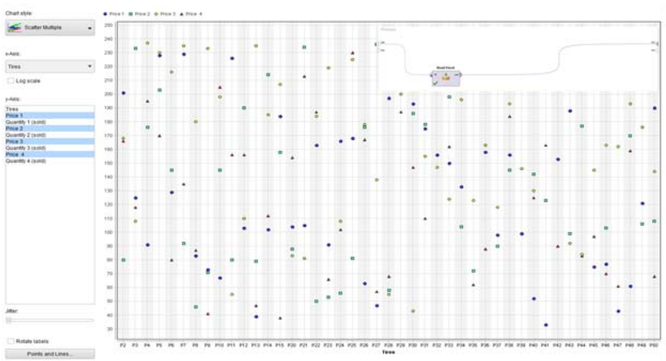

Fig. 14. Scatter Multiple plot of Rapid Miner tool. Inset: "Read Excel" object and related workflow.

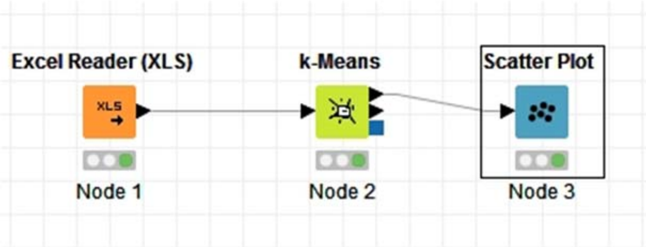

Fig. 15. KNIME Workflow: k-Means algorithm.

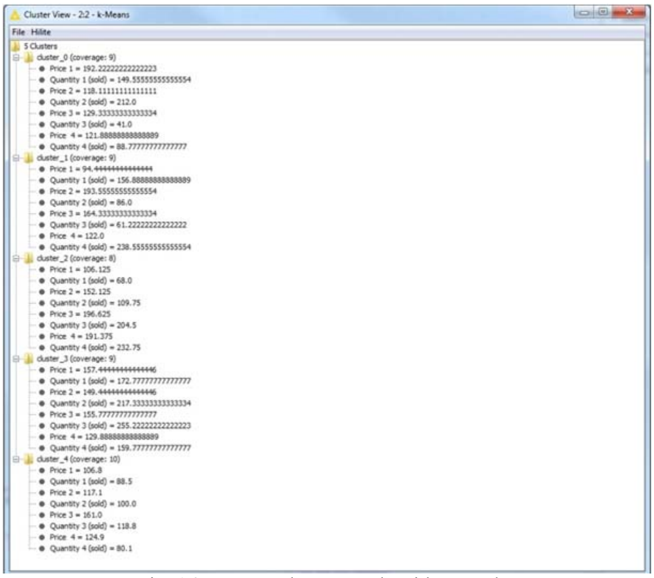

Fig. 16. KNIME: k-Means algorithm results

We focus now our attention on the key interpretation of results of Fig. 16 and Fig. 17, by considering a single tire model. For example the tire model P2 belongs to the cluster 0 (see first element of the plot of Fig. 17) which takes into account the 4 competitors prices and the sold quantities. Observing the data reported in Fig. 16 about the maximum quantity (quantity 2 parameter corresponding to 212), the best price to adopt for the tire P2 is provided by the parameter Price 2 (118.11 Euro). This is the decision support system to establish the best price by considering the high sold quantities and the competitor prices.

By considering tyres model P3 we observe that it belongs to cluster 1; the best price to adopt by observing high sold quantities (Quantity 4 =238.5) is Price 4 (122). This analysis procedure can be applied for each model type. We observe that the horizontal axis of Fig. 17 represents the tires types (Model P2, P3, P4, etc.), besides the vertical axis of Fig. 17 indicates the clusters (cluster 0, cluster 1, cluster 2, cluster 3, cluster 4)

Fig. 17. KNIME: clustering results.

In Fig. 18 we illustrates the clustering results in a table format , the lower and upper bound limits of cluters, and all clustering results. The data are viewed by the selecting the object of the node 2 of Fig. 15.

Row ID	D Price 1	D Quantity 1 (sold)	D Price 2	D Quantity 2 (sold)	D Price 3	D Quantity 3 (sold)	D Price 4	D Quantity 4 (sold)
cluster_0	192.222	149.556	118.111	212	129.333	41	121.889	88.778
cluster_1	94.444	156.889	193.556	86	164.333	61.222	122	238.556
cluster_2	106.125	68	152.125	109.75	196.625	204.5	191.375	232.75
cluster_3	157.444	172.778	149.444	217.333	155.778	255.222	129.889	159.778
cluster_4	106.8	88.5	117.1	100	161	118.8	124.9	80.1

Columns: 8	Column Type	Column Index	...	...	...	...	Lower Bound	Upper Bound
Price 1	Number (double)	0					94.444	192.222
Quantity 1 (sold)	Number (double)	1					68	172.778
Price 2	Number (double)	2					117.1	193.556
Quantity 2 (sold)	Number (double)	3					86	217.333
Price 3	Number (double)	4					129.333	196.625
Quantity 3 (sold)	Number (double)	5					41	255.222
Price 4	Number (double)	6					121.889	191.375
Quantity 4 (sold)	Number (double)	7					80.1	238.556

Row ID	S Tires	I Price 1	I Quantit...	I Price 2	I Quantit...	I Price 3	I Quantit...	I Price 4	I Quantit...	S Cluster
Row0	P2	201	160	80	245	168	65	166	163	cluster_0
Row1	P3	125	120	233	76	108	112	118	166	cluster_1
Row2	P4	91	20	176	225	237	203	195	258	cluster_2
Row3	P5	228	90	203	246	230	242	170	103	cluster_3
Row4	P6	129	71	145	135	216	141	80	80	cluster_4
Row5	P7	229	163	92	253	235	22	135	6	cluster_0
Row6	P8	83	293	46	216	180	125	87	18	cluster_0
Row7	P9	73	157	71	270	233	240	41	176	cluster_3
Row8	P10	67	293	145	106	198	2	205	265	cluster_1
Row9	P11	226	231	80	206	55	111	156	84	cluster_0
Row10	P12	103	63	190	90	110	59	156	68	cluster_4
Row11	P13	39	170	79	104	235	196	47	128	cluster_4
Row12	P14	102	119	214	206	185	113	112	48	cluster_4
Row13	P15	184	109	158	190	207	265	38	147	cluster_3
Row14	P20	104	89	88	136	83	74	154	111	cluster_4
Row15	P21	105	226	234	270	81	275	213	213	cluster_3
Row16	P22	163	29	50	9	184	173	187	49	cluster_4
Row17	P23	91	133	53	58	219	46	66	239	cluster_1
Row18	P24	166	123	56	117	108	123	102	19	cluster_4
Row19	P25	168	98	81	104	225	181	230	212	cluster_2
Row20	P26	63	194	176	198	178	226	167	297	cluster_2
Row21	P27	47	68	236	39	138	67	57	202	cluster_1
Row22	P28	197	18	58	266	55	5	68	19	cluster_0
Row23	P29	204	72	236	185	200	31	187	138	cluster_0
Row24	P30	193	149	186	149	43	4	147	101	cluster_0
Row25	P31	175	230	178	179	155	217	110	45	cluster_3
Row26	P32	156	4	240	13	147	212	225	244	cluster_2
Row27	P33	150	161	198	17	124	96	162	244	cluster_1
Row28	P34	133	36	104	93	196	9	216	136	cluster_4
Row29	P35	208	226	72	111	123	266	62	231	cluster_3
Row30	P36	158	210	240	237	163	280	88	268	cluster_3
Row31	P37	98	204	90	210	118	290	219	53	cluster_3
Row32	P38	156	10	145	84	193	290	184	125	cluster_2
Row33	P39	99	67	234	26	146	65	242	265	cluster_1
Row34	P40	52	155	142	81	130	40	125	56	cluster_4
Row35	P41	33	24	123	118	224	77	163	283	cluster_2
Row36	P42	153	264	210	232	239	21	90	297	cluster_1
Row37	P43	188	103	99	243	92	222	228	202	cluster_3
Row38	P44	207	47	177	294	84	4	83	215	cluster_0
Row39	P45	75	53	216	188	145	46	97	172	cluster_1
Row40	P46	77	30	103	29	163	260	70	106	cluster_4
Row41	P47	43	253	217	32	162	96	61	297	cluster_1
Row42	P48	61	175	170	76	193	194	159	275	cluster_2
Row43	P49	121	19	106	60	176	253	208	168	cluster_2
Row44	P50	190	213	108	94	144	2	68	55	cluster_0

Fig. 18. KNIME: other clustering results.

IV. Conclusion

The goal of this paper was to provide results of a an industrial project oriented on intelligent dynamic tires price definition. The project includes smart mobile technologies and data mining processing able to improve a decision support system concerning tires price to adopt. In the paper were presented all the design steps, some tests validating the software implementation and clustering results. The best tires price were selected dynamically by observing during the time together sold quantities and competitors price. The data mining engine could be applied also for the sales and price predictions by implementing a linear regression algorithm [17] or Artificial Neural Networks – ANN- workflow based on time series forecasting approach [18].

Acknowledgment

The work has been developed in the frameworks of the Italian projects: “Modello dinamico di identificazione della tipologia e del prezzo ottimale dei pneumatici mediante applicazione web (Dynamic model for identifying the type and optimal price of tires through a web application): “*Tires Detection and Dynamic Pricing*”. Authors gratefully thanks Vito Custodero and Elisabetta Valenzano for their support in the realization of this work. special thanks are addressed to William Ciavarella for his knowledge in tires marketing.

REFERENCES

[1] Y. Narahari, C.V.L. Raju, K. Ravikumar, S. Shah, "Dynamic Pricing Models for Electronic Business," Springer, Sādhanā, April 2005, Volume 30, Issue 2, pp 231–256.

[2] C. V. L. Raju · Y. Narahari · K. Ravikumar, "Learning dynamic prices in electronic retail markets with customer segmentation," Ann Oper Res (2006) 143: 59–75.

[3] A. Ghose, T. Mukhopadhyay, U. Rajan, V. Choudhary, "Dynamic pricing: a strategic advantage for electronic retailers," 2002 — Twenty-Third International Conference on Information Systems, pp. 305-315.

[4] M. Kung, K. B. Monroe, J. L. Cox, "Pricing on the internet," Journal of Product & Brandmanagement, vol. 11, no. 5, 2002 , pp. 274-287.

[5] A. V. den Boer, "Dynamic Pricing and Learning: Historical Origins, Current Research, and New Directions," (December 2013). ISSN 1874−4850.

[6] D. Bergemann, J. Valimaki, "Dynamic pricing of new experience goods," Journal of Political Economy, 2006, vol. 114, no. 4.

[7] Karakaya, Fahri and D. Steven White (2000). E-Commerce and Marketing Strategy, Academy of Business Administration, 353-358.

[8] Lal, Rajiv and Miklos Sarvary (1999). When and How is the Internet Likely to Decrease Price Competition, Marketing Science, 18(4), 485-503.

[9] Suri, Rajneesh, Mary Long, and Kent B. Monroe (2001). The Impact of the Internet and Consumer Motivation on the Evaluation of Prices, Journal of Business Research, 55, 1-12.

[10] C. Patel, D. Shah, A. Patel, "Automatic Number Plate Recognition System (ANPR): A Survey," International Journal of Computer Applications (0975 – 8887) Volume 69– No.9, May 2013.

[11] Anton Satria Prabuwono and Ariff Idris, "A Study of Car Park Control System Using Optical Character Recognition ," in International Conference on Computer and Electrical Engineering, 2008, pp. 866-870.

[12] Christos Nikolaos E. Anagnostopoulos, Ioannis E. Anagnostopoulos, Vassili Loumos, and Eleftherios Kayafas, "A License Plate-Recognition Algorithm for Intelligent Transportation System Applications," pp. 377- 392, 2006.

[13] Prathamesh Kulkarni, Ashish Khatri, Prateek Banga, Kushal Shah, "Automatic Number Plate Recognition (ANPR) System for Indian conditions," IEEE 19th International Conference Radioelektronika, 2009. RADIOELEKTRONIKA '09, 22-23 April 2009.

[14] Pratiksha Jain, "Automatic License Plate Recognition using OpenCV," International Journal of Computer Applications Technology and Research, Volume 3– Issue 12, 756 - 761, 2014, ISSN:- 2319–8656.

[15] Alessandro Massaro, Angelo Galiano, Donato Barbuzzi, Leonardo Pellicani, Giuseppe Birardi, Davide Donato Romagno, Luisa Frulli, "Joint Activities of Market Basket Analysis and Product Facing for Business Intelligence oriented on Global Distribution Market: examples of data mining applications," International Journal of Computer Science and Information Technologies, Vol. 8 (2) , 2017, 178-183.

[16] Vijay Kotu, Bala Deshpande, "Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner," Executive Editor: Steven Elliot, Elsevier 2015.

[17] A. Massaro, D. Barbuzzi, V. Vitti, A. Galiano, M. Aruci, G. Pirlo, "Predictive Sales Analysis According to the Effect of Weather," Proceeding of the 2nd International Conference on Recent Trends and Applications in Computer Science and Information Technology, Tirana, Albania, November 18 - 19, pp. 53-55, 2016.

[18] Mehdi Khashei, and Mehdi Bijari, "An artificial neural network (p, d, q) model for timeseries forecasting," Expert Systems with Applications, Vol. 37, Issue 1, January 2010, Pages 479-489.