

## **A Comparative Study of Machine Learning Methods for Automatic Classification of Academic and Vocational Guidance Questions**

<https://doi.org/10.3991/ijim.v14i08.13005>

Omar Zahour <sup>(✉)</sup>, El Habib Benlahmar, Ahmed Eddaouim, Oumaima Hourrane  
Hassan II University, Casablanca, Morocco  
orzahour@gmail.com

**Abstract**—Academic and vocational guidance is a particularly important issue today, as it strongly determines the chances of successful integration into the labor market, which has become increasingly difficult. Families have understood this because they are interested, often with concern, in the orientation of their child. In this context, it is very important to consider the interests, trades, skills, and personality of each student to make the right decision and build a strong career path. This paper deals with the problematic of educational and vocational guidance by providing a comparative study of the results of four machine-learning algorithms. The algorithms we used are for the automatic classification of school orientation questions and four categories based on John L. Holland's Theory of RIASEC typology. The results of this study show that neural networks work better than the other three algorithms in terms of the automatic classification of these questions. In this sense, our model allows us to automatically generate questions in this domain. This model can serve practitioners and researchers in E-Orientation for further research because the algorithms give us good results.

**Keywords**—Academic and vocational guidance, E-orientation, Machine learning, Automatic classification, Comparative study.

### **1 Introduction**

The classification of questions is a problem that has already been studied by several researchers in this field, but most of the work is domain-specific or limited to a high-level classification.

Anbuselvan and R.Ahmed [1] proposed an SVM-based method for the same task. The question is first analyzed and numbered, the parts of the speech are labeled, the stop words are deleted, the data is truncated and many features are extracted. The feature selection steps are performed prior to transmitting the data to a carrier vector machine for training. The same treatment is also done for test questions, which can take a long time to get results in real-time.

Marco Pota [2] propose a feature-based method, in which features related to a subset of questions such as keywords, how - all / some words, leading verbs and various other such features were extracted from the texts a classifier.

For Natural Language Processing (NLP) Convolutional neural networks (CNNs) have already been used in some works. Collobert and J.Weston [3] first proposed the idea of a convolutional neural network architecture, which includes lookup tables and hyperbolic hard tangents. Kalchbrenner and P.Blunson [4] proposed a simplified version of Collobert's network, which was used to classify Twitter's questions and opinions. They used the concept of k-max pooling. Yoon Kim [5] developed Kalchbrenner's work to add various machine-learning strategies, such as regularization, to improve network performance.

For the time, the question classification has mainly been studied in the context of open-domain TREC (Text REtrieval Conference) questions [6], with smaller recent datasets available in biomedical [7] [8] and education [9]. The TREC corpus of questions from the open-domain is a set of questions associated with a taxonomy developed by Li and Roth [10] that includes 6 types of coarse responses (such as entities, locations, and numbers) and 50 fine-grained types (for example, specific types of entities, such as animals or vehicles). While a wide variety of syntactic, semantic and other features and classification methods have been applied to this task, culminating in an almost perfect classification performance [11], recent work has shown that QC methods developed on TREC issues usually fail to transfer to datasets with more complex issues such as those in the biomedical field [7], probably due in part to the simplicity and syntactic regularity of questions and the possibility of simpler term frequency models achieve near-ceiling performance [12].

In this world, the educational and guidance system of each country seeks to help the students or the laureates of higher education institutions and vocational training institutes to make their choice.

According to Ali Boulahcen [13] and through his analysis, he noticed that there is no real process of educational guidance in Morocco, but there is only a summary process in the context, within a few seconds, one decides on the fate of the pupil that based solely on his academic value then translated by a numerical note.

This means that the Moroccan school institution is based at least on selection criteria and not on orientation [13]. In this context, our goal is to set up an E-Orientation system that is interested in the automation of the orientation task, thanks to the evolution of information technologies. The realization of this electronic guidance system requires the classification then modeling and integration of user preferences in this system. In this paper, we used the Multi-Class Neural Networks algorithm to classify the different questions according to John L. Holland's RIASEC topology.

This document is organized as follows: Section 2 is devoted to a review of the literature of different theories of educational and vocational guidance, including the theory of John L. Holland. Section 3 is devoted to the various algorithms for the automatic classification of text that we will use in our model. Section 4 deals with the experimental evaluation of each classification algorithm with the results obtained. Finally, section 5 covers the conclusion with research perspectives.

## **2 Related Work**

The guiding approach is based on theories and studies related to career choice and career development. These include Hoyt's concept of career education, Gardner's theory of multiple intelligences and Holland's typology of professional interests [14]. Holland's theory of vocational choice (1997) [15], is the result of the work of American psychologist and researcher "John Holland (1919-2008)". The results of his research argue that their skills, interests, and personality would determine the association of workers to one type of career.

Some activities would be better suited to one type of person than another would. It constitutes the theoretical anchoring of our classification model and serves as a basis for many psychometric tools, including the Hexa 3d professional interest's questionnaire. Although this theory, dating from the mid-1960s is still widely used [16] and has been the subject of numerous studies [17], [18].

To briefly explain his theory, Holland (1997) [15] formulates several hypotheses according to professional interests that are a mode of expression of personality. Therefore, he considers the choices of orientation as a mode of expression of this personality and distinguishes six types of personality (RIASEC), according to aptitudes, personality traits, values, and beliefs.

Of all the models related to career development, the Holland model has been the subject of the greatest number of analyzes and studies.[19]. Among those conducted on the structure of interests across gender and ethnic populations, a number demonstrates the consistency of the arrangement of types and their proximity on a hexagonal and spherical model [18], [20], [21]. This debate focuses more on the geometric regularity of the hexagon and on the correspondence distances between the different types. Vrignaud and Bernaud (1994) validated other things such as the structure of the Holland model in France [22].

Professional activities, as well as work environments, tend to bring together people who share common interests to a certain extent. The choice of a profession or trade is a form of expression of the personality of an individual; it is the theory of vocational interests. As well as, the person-work environment combination is the most widely used method in the world of educational and vocational guidance.

The theory of vocational choice distinguishes six categories of professional interest (realistic, investigative, artistic, social, enterprising, and conventional) corresponding to different personality profiles. Holland represents them according to a hexagonal model illustrated in Fig.1 [23].

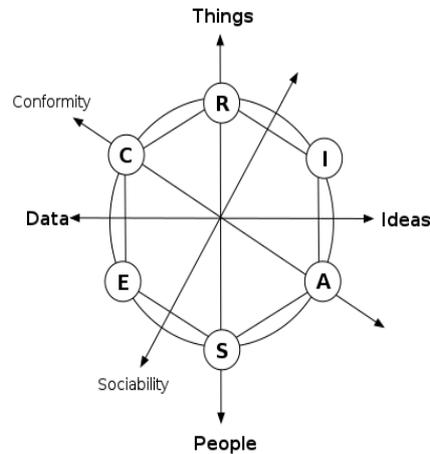


Fig. 1. Representation of Holland's circular Model (RIASEC)

According to Holland's theory and previous research, they have confirmed the profession or trade chosen by a person which is a form of expression of his personality. Therefore, it is related to the type to which he belongs.

The affiliation of a worker to one of the six types would be determined by his aptitudes, by certain traits of his personality and interests. So, according to Holland, people of the same type would be attracted to the same kind of work. Why? Because these people are similar in their personality, in the fact they pursue similar objectives and have the same physical or psychological dispositions with regard to their work. All persons can be divided into six professional types.

The typology of a person is established by measuring his degree of affinity with each of the six types, to place them in order of importance, of the type that corresponds most to him. For most people, it is mostly the first two or three types of personal classification that determine their way of being and acting in their personal and professional lives. For example, a person whose dominant type is "Investigator" and who has affinities with the "Realist" type; we will say that he has an "IR" profile. To further characterize this person's typology, it is possible to consider the third type which it most closely resembles and to say the case where it is of the "Social" type and is this person has an "IRS" profile?

These types can be combined in all sorts of ways and their combination determines the personality.

- **The Realistic type:** People of this type take pleasure in carrying out concrete tasks. Adroit with their hands, they know how to coordinate their actions. They are happy to use tools, are adept at appliances, machines, vehicles. No problem to tinker or repair what is down. Realists often have a sense of mechanics and precision. Many practice their profession outdoors rather than indoors. Their work often requires good physical stamina and even athletic abilities.

- **The Investigator type:** Most investigators are not afraid of "theory", on the contrary. They like to collect data, make assumptions, look for solutions to solve problems as we do in maths. The "investigators" take the time of the observation; they are often "secondary" unlike the impulsiveness that acts without taking the time of the analysis. So, they like to be absorbed in their thoughts, play with ideas. In the work, we appreciate their intellectual rigor and their sense of method, but as a team, their character may seem a bit cold and distant.
- **The Artistic type:** Artist profiles are interested in creative work, be it visual art, literature, music, advertising or theater. Independent and non-conformist, they are comfortable in situations that are out of the ordinary. They are endowed with great sensitivity and imagination. Although they are discouraged by methodical and routine tasks, they are nevertheless able to work with discipline to perfect their artistic talent and to carry out long-term work.
- **The Social type:** People of this type like to be in contact with others in order to help them, to inform them, to educate them, to entertain them, to treat them or to promote their growth. They are interested in human behaviors and are concerned about the quality of their relationships with others. They use their knowledge and their feelings and emotions to act and interact.
- **The Entrepreneurial type:** People of this type like to influence their surroundings. Their decision-making ability, sense of organization and a particular ability to communicate their enthusiasm support them in their goals. They know how to sell ideas as much as material goods. They have a sense of organization, planning, and initiative and know-how to carry out their projects. They know how to be bold and efficient.
- **The Conventional type:** People of this type have a preference for specific, methodical activities that focus on a predictable outcome. They are concerned about the order and the good material organization of their environment. They prefer to conform to well-established conventions and clear instructions rather than to act with improvisation. They like to calculate, classify, maintain registers. or folders. They are effective in any job that requires accuracy and ease in routine tasks.[24].

### 3 Materials and Methods

Classification systems for the best-performing questions tend to use a rule-based custom template matching [25] [11], or a combination of basic learning approaches. of rules and machine learning [26], to the detriment of model construction time.

Recent research on the methods learned has shown that a large number of CNN variants [27] and LSTM [12] achieve similar precision on the TREC question classification; these models presenting at best small gains compared to simple models. Term frequency models. These recent developments echo the observations of Roberts and M.Fizman [7], who have shown that existing methods beyond term frequency models fail to generalize to questions in the medical field.

In the education sector, researchers Godea. A and Nielsen.R [9] collected 1,155 questions in class and classified them into 16 categories. To allow a detailed study of the classification of questions in the scientific field.

The process of classifying a text collection is to label each text with one or more predefined classes (Categories). In this process, an algorithm is first designed then it is driven with a set of specific characteristics, for example, word occurrences or even theme distributions in a document. Once trained, the algorithm is used to label new texts, but these are different from the texts used during training. The algorithm is evaluated on the number of classification errors obtained during the learning phase and during the test phase.

When we are training the classification algorithm, the extraction phase of the characteristics is used for learning crucial. These Characteristics extracted from texts that are typically derived from a large vector space. This space is constructed with vector modeling of words using distributional semantics [28].

Data science or statistical algorithms are further classified into multiple machines learning specific algorithmic categories:

- Supervised learning algorithms (label and output known).
- Unsupervised learning algorithms (label and output not known).
- Reinforced learning algorithms (reward-based agent action).
- Semi-supervised learning algorithms (mix of supervised and unsupervised).

These algorithms, in turn, contain multiple sub-algorithms and types (see Table I). For example, a few algorithms fall under the category of parametric, whereas others are non-parametric. In parametric algorithms, information about the population is completely known which not the case with non-parametric algorithms is. Typically, parametric models deal with a finite number of parameters, whereas non-parametric learning models are capable of dealing with an infinite number of parameters. Therefore, the training data grows the complexity of nonparametric models increases. Linear regression, logistic regression, and Support vector machines are examples of parametric algorithms. K-nearest neighbor and decision trees are non-parametric learning algorithms. These algorithms are computationally faster in comparison to their nonparametric companions. As TABLE 1 depicts, the machine learning algorithms are large in number [29].

**Table 1.** Machine learning Algorithmus

<b>Supervised Learning</b>	<b>Unsupervised Learning</b>	<b>Reinforcement Learning</b>
Artificial neural network	Artificial neural network	Q-learning
Bayesian statistics	Association rule learning	Learning automata
Case-based reasoning	Hierarchical clustering	
Decision trees	Partitioned clustering	
Learning automata		
Instance-based learning		
Regression analysis		
Linear classifiers		
Decision trees		
Bayesian networks		
Hidden Markov models		

In this section, we will describe the different classification algorithms used in our research

### 3.1 Multiclass decision forest

The decision forest algorithm is an ensemble learning method for classification. The algorithm works by creating several decision trees and then voting on the most popular output class. Voting is a form of aggregation, in which each tree in a classification decision forest generates a non-standard frequency histogram of labels. The aggregation process adds these histograms and normalizes the result to obtain the "probabilities" for each label. Trees that have high confidence in the forecasts have a greater weight in the final decision of the set.

Decision trees, in general, are non-parametric models, which means that they support data with varied distributions. In each tree, a simple test sequence is executed for each class, increasing the levels of a tree structure until a leaf node (decision) is reached.

Decision trees have many advantages, they can represent non-linear decision limits, they are effective in calculating and using memory during training and prediction, and they perform an integrated selection and classification of features are resistant in the presence of noise characteristics.

The decision forest classifier in Azure Machine Learning Studio (Classic) consists of a set of decision trees. In general, ensemble models provide better coverage and accuracy than single decision trees.

### 3.2 Multiclass decision jungle

Decision Jungles are a recent extension of Decision Forests. A decision jungle consists of a set of decision-directed acyclic graphs (DAGs). The decision jungles have the following advantages; By allowing tree branches to merge, a decision DAG generally has a smaller memory footprint and better generalization performance than a decision tree, but at the cost of slightly longer training time. Additionally, decision

jungles are non-parametric models, which can represent nonlinear decision boundaries. Finally, they perform an integrated selection and classification of features and are resistant to noisy features.

### 3.3 Multiclass regression logistic

Logistic regression classification is a supervised learning method and therefore requires a tagged dataset. You train the model by providing the model and dataset labeled as input to a module such as the Train Model or Tune Model Hyperparameters. The driven model can then be used to predict the values of new input examples.

Logistic regression is a well-known method in statistics that is used to predict the probability of a result and is particularly popular for classification tasks. The algorithm predicts the probability of occurrence of an event by adjusting the data to a logistic function. For more details on this implementation, see the Technical Notes section. In multi-class logistic regression, the classifier can be used to predict multiple outcomes.

Multinomial logistic regression is a form of logistic regression, which used to predict a target variable; it has more than 2 classes. It is a modification of logistic regression using the softmax function instead of the sigmoid function, and the cross-entropy loss function. The softmax function squashes all values to the range [0, 1] and the sum of the elements is one.

$$\text{soft max}(x)_i = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \quad (1)$$

Cross entropy is a measure of how different 2 probability distributions are near to each other. If p and q are discrete, we have:

$$H(p, q) = -\sum_x p(x) \text{Log } q(x) \quad (2)$$

This function has a range of [0, inf], it is equal to 0 when p=q and infinity then p is very small compared to q or vice versa. For example x, the class scores are given by vector  $z=Wx+b$ , where W is a  $C \times M$  matrix and b is a length C vector of biases. We define the label y as a one-hot vector equal to 1 for the correct class c and 0 everywhere else. The loss for a training example x with predicted class distribution y and correct class c will be:

$$\hat{y} = \text{softmax}(z) \quad (3)$$

$$\begin{aligned}
 loss &= H(y, \hat{y}) \\
 &= -\sum_i y_i \text{Log } \hat{y}_i \\
 &= -\text{Log } \hat{y}_c
 \end{aligned}
 \tag{4}$$

As in the binary case, the loss value is exactly the negative log probability of a single example  $x$  having true class label  $c$ . Thus, minimizing the sum of the loss over our training example is equivalent to maximizing the log-likelihood. We can learn the model parameters  $W$  and  $b$  by performing gradient descent on the loss function with respect to these parameters. There are two common methods to perform multi-class classification using the binary classification logistic regression algorithm: one-vs-all and one-vs-one. In one-vs-all, we train  $C$  separate binary classifier for each class and run all those classifiers on any new example  $x$ , we want to predict and take the class with the maximum score. In one-vs-one, we train  $C$  to choose 2 classifiers =  $C(C-1)/2$  one for each possible pair of class and choose the class with maximum votes while predicting for a new example.

### 3.4 Multiclass neural network

A neural network is a set of interconnected layers. The inputs are the first layer and are connected to an output layer by an acyclic graph comprised of weighted edges and nodes.

Between the input and output layers, you can insert multiple hidden layers. Most predictive tasks can be accomplished easily with only one or a few hidden layers. However, recent research has shown that deep neural networks (DNN) with many layers can be very effective in complex tasks such as image or speech recognition. The successive layers are used to model increasing levels of semantic depth.

The relationship between inputs and outputs is learned from training the neural network on the input data. The direction of the graph proceeds from the inputs through the hidden layer and to the output layer. All nodes in a layer are connected by the weighted edges to nodes in the next layer.

To compute the output of the network for a particular input, a value is calculated at each node in the hidden layers and in the output layer. The value is set by calculating the weighted sum of the values of the nodes from the previous layer. An activation function is then applied to that weighted sum. For example, neural networks of this type can be used in complex computer vision tasks, such as recognition of numbers or letters, document classification, and pattern recognition.

Classification using neural networks is a supervised learning method and therefore requires a tagged data set that includes a label column. You can train the model by providing the tagged model and dataset as input for Train Model or Tune Model Hyperparameters. The driven model can then be used to predict the values of the new input examples.

A neural network is a set of interconnected layers. The inputs are the first layer and are connected to an output layer by an acyclic graph composed of weighted edges and nodes. We can insert multiple hidden layers between the input and output layers. Most predictive tasks can be accomplished easily with one or more hidden layers. However, Deep Neural Networks (DNNs) [30], [31] with many layers can be very effective for complex tasks such as image recognition or speech. Successive layers are used to model increasing levels of semantic depth. The relationship between inputs and outputs is learned during the formation of the neural network on the input data. The chart direction passes inputs to the hidden layer and the output layer. All the nodes of a layer are connected by the weighted edges to the nodes of the next layer.

To calculate the network output for a particular input, a value is calculated at each node of the masked layers and the output layer. The value is defined by calculating the weighted sum of the values of the nodes of the previous layer. An activation function is then applied to this weighted sum.

We use a multiclass neural network module to predict a multi-valued target knowing that neural networks of this type could be used in complex computer vision tasks, such as recognition of numbers or letters, classification of documents, of text (Questions) and for pattern recognition. In this sense classification, using neural networks is a supervised learning method. It, therefore, requires a tagged data set comprising a label column.

## **4 Proposed Method**

Our proposed system is based on the four algorithms described in the second part of this article that follows supervised learning. The goal is to discover an underlying structure of the data. This algorithm requires a tagged data set. The E-Orientation Data Orientation Data Set is divided into two series, such as training data and test data. The classification performed by the algorithm used in our model is based on the knowledge acquired by the learning data during the learning process.

Our dataset was collected from the RIASEC test based on Holland's theory [32], [33], [34], it contains two columns namely:

Question: It contains questions and statements that measure either the occupations or the activities or abilities or the personality of the users.

Categories: we have four classes (labels) of categories namely:

1. Activity.
2. Occupations.
3. Abilities.
4. Personality.

In our research work on Guidance Classification, we used the Azure Machine Learning Studio [35] tool which is a collaborative drag-and-drop tool that we can use to create, test, and deploy predictive analytics solutions on our data. Machine Learning Studio publishes templates as a web of services that can be easily consumed

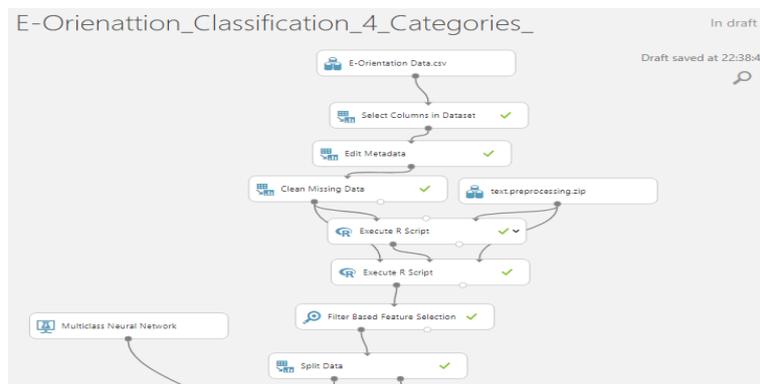
by custom applications. Machine Learning Studio is the meeting place of data science, predictive analytics, cloud resources, and our data.

## 5 Experiment and Results

The experimental steps described and illustrated in Fig.2. They are explained below:

- a) **Importing the dataset:** We import our dataset entitled "E-Orientation Data" that we collected from several websites from our local disk on Azure ML Studio to be used for the experiment and Category names that we have been used as a class tag or attribute to predict.
- b) **Preprocessing and preparing the dataset:** The dummy column headers have been replaced by meaningful column names by using the metadata editor. In addition, missing values have been cleared by deleting the entire line containing the missing value.
- c) **Feature engineering:** After the processing phase of the dataset, we will use the feature hashing module to convert the raw text of the questions into integers; and use the integer values as input entities of the model. Figure 3 represents our model.
- d) **Split the data and parameter settings:** We have divided the data of "E-Orientation Data" as 70% of the data for training and 30% for the test. Then for the Multiclass Neural Networks algorithm, we applied it with the default settings for model formation. The parameters have been set by using the "Tune model hyperparameters".
- e) **The model:** each time We used one of these four algorithms
- f) **Score and evaluate the model:** The Evaluate model visualizes the results through the confusion matrix.

For the schema of our model, we can summarize it in the following figure knowing that for each algorithm we keep the same steps described in the figure except that we change the algorithm used.



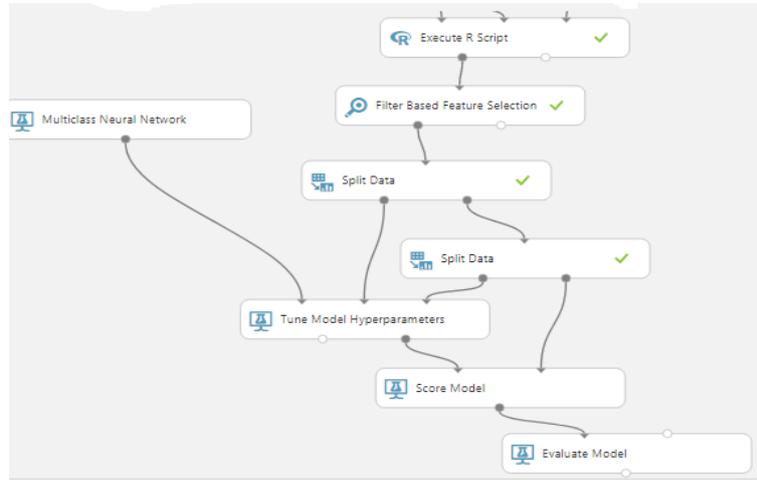


Fig. 2. Schema of model

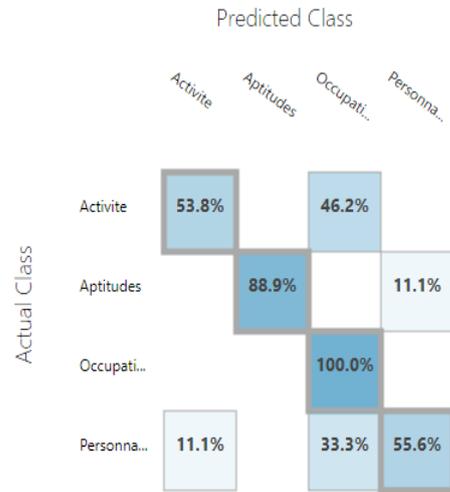
Table 2. The metric values of the four Algorithms

Method	Overall accuracy	Average accuracy	Micro-averaged precision	Macro-averaged precision	Micro-averaged recall	Macro-averaged recall
Multiclass Decision Forest	0,75	0,875	0,75	0,824811	0,75	0,745726
Multiclass decision Jungle	0,75	0,875	0,75	0,824811	0,75	0,745726
Multiclass regression Logistic	0,795455	0,897727	0,795455	0,845833	0,795455	0,784188
Multiclass neural network	0,818182	0,909091	0,818182	0,8875	0,818182	0,811966

According to the results shown in this last table, we note that the results obtained by the Multiclass neural network algorithm are the best followed by the results of the Multiclass Regression Logistic algorithm and for the two algorithms we see that the results are the same., this shows us that the best algorithm to use is the Multiclass neural network algorithm [36].

For the matrix of confusion concerning the algorithm Multiclass decision forest, we have obtained the following in figure number 02.

E-Orientattion\_Classification\_4\_Categories\_ > Evaluate Model > Evaluation results

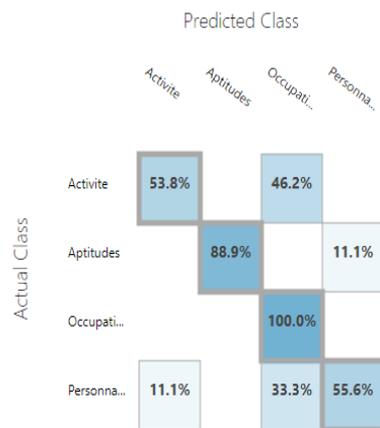


**Fig. 3.** Confusion Matrix of Multiclass Decision Forest

For the matrix of confusion concerning the algorithm Multiclass Decision Jungle, we have obtained the following in figure number 03.

E-Orientattion\_Classification\_4\_Categories\_ > Evaluate Model > Evaluation results

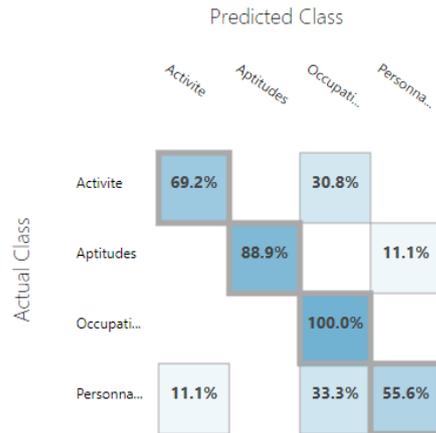
[Confusion Matrix](#)



**Fig. 4.** Confusion Matrix of Multiclass Decision Jungle

For the matrix of confusion concerning the algorithm Multiclass Regression Logistic we have obtained the following in the figure number 04

E-Orientattion\_Classification\_4\_Categories\_ > Evaluate Model > Evaluation results

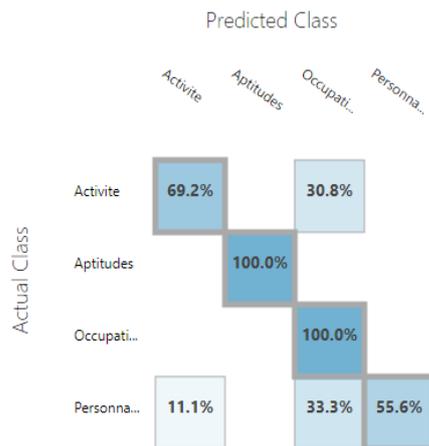


**Fig. 5.** Confusion Matrix of Multiclass Regression Logistic

For the matrix of confusion concerning the algorithm Multiclass Neural Network, we have obtained the following in figure number 05

E-Orientattion\_Classification\_4\_Categories\_ > Evaluate Model > Evaluation results

▾ Confusion Matrix



**Fig. 6.** Confusion Matrix of Multiclass Neural Network

## 6 Conclusion

In this article, we defined and applied the four machine learning algorithms used for text classification. We conclude that multi-class neural networks work better than the other three machine learning algorithms.

The Multiclass Neural Network algorithm used in our classification model of Academic and Professional Class Orientation Issues is implemented using Azure Machine Learning Studio. In fact, we found that the supervised method gives very good precision. This method can also be used to automatically generate academic and vocational orientation questionnaires by knowing the class of these proposed new questions in advance, and we can view this research question as a perspective. This automatic classification model using machine-learning algorithms can also help E-guidance researchers in the development process in this area.

As future work, we focus on the use of social network analysis, for example, using Twitter's sentiment analysis as a feature to determine the class of questions and interests of students and faculties of educational institutions. 'Education. The emergence of a new multi-label classification approach called BERT [37], the acronym for Bidirectional Encoder Representations from Transformers, is a language model (in) developed by Google in 2018. This method has significantly improved automatic language processing algorithms; the application of this method in our next work is an issue in order to compare the results obtained by the latter method with the results obtained by these four algorithms used in this research work. to develop a system of E-orientation is our goal knowing that online services (evaluation, learning) have shown their great effectiveness according to several researchers [38][39][40][41].

## 7 References

- [1] A. Sangodiah, R. Ahmad, and W. F. W. Ahmad, "A review in feature extraction approach in question classification using Support Vector Machine," in Proceedings - 4th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2014, 2014. <https://doi.org/10.1109/iccsce.2014.7072776>
- [2] G. D. Pietro Marco Pota, Angela Fuggi, Massimo Esposito, "Extracting Compact Sets of Features for Question Classification in Cognitive Systems: A Comparative Study," in 2015 ,10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, pp. 551–55. <https://doi.org/10.1109/3pgcic.2015.118>
- [3] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," J. Mach. Learn. Res., 2011.
- [4] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," in 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference, 2014. <https://doi.org/10.3115/v1/p14-1062>
- [5] Y. Kim, "Convolutional neural networks for sentence classification," in EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, 2014. <https://doi.org/10.3115/v1/d14-1181>

- [6] E. M. Voorhees and D. M. Tice, "Building a question answering test collection," SIGIR Forum (ACM Spec. Interes. Gr. Inf. Retrieval), 2000. <https://doi.org/10.1145/345508.345577>
- [7] K. Roberts, H. Kilicoglu, M. Fiszman, and D. Demner-Fushman, "Automatically classifying question types for consumer health questions," AMIA Annu. Symp. Proc., 2014. <https://doi.org/10.3115/v1/w14-3405>
- [8] M. Wasim, M. N. Asim, M. U. Ghani Khan, and W. Mahmood, "Multi-label biomedical question classification for lexical answer type prediction," J. Biomed. Inform., 2019. <https://doi.org/10.1016/j.jbi.2019.103143>
- [9] A. Godea and R. Nielsen, "Annotating educational questions for student response analysis," in LREC 2018 - 11th International Conference on Language Resources and Evaluation, 2019.
- [10] X. Li and D. Roth, "Learning question classifiers," 2002.
- [11] H. T. Madabushi and M. Lee, "High accuracy rule-based question classification using question syntax and semantics," in COLING 2016 - 26th International Conference on Computational Linguistics, Proceedings of COLING 2016: Technical Papers, 2016.
- [12] W. Xia, W. Zhu, B. Liao, M. Chen, L. Cai, and L. Huang, "Novel architecture for long short-term memory used in question classification," Neurocomputing, 2018. <https://doi.org/10.1016/j.neucom.2018.03.020>
- [13] A. Boulahcen, "Le processus d'orientation scolaire au MarocA Sociological Analysis of the Careers Advice Process in Moroccan SchoolsUn análisis sociológico del proceso de orientación escolar en Marruecos," Rev. Int. d'éducation Sèvres, 2005. <https://doi.org/10.4000/ries.1427>
- [14] G. & M. D. Dupont.P, "Guide for Educational and Vocational Information and Guidance: Elementary and Secondary Cycle Three, Youth Sector," Prov. Support Gr. Youth Orientat. approach to Sch. Univ. Sherbrook, Sherbrook, Canada., 2002.
- [15] J. L. Holland, "Making vocational choices: A theory of vocational personalities and work environments," Englewood Cliffs NJ PrenticeHall, 1997.
- [16] M. M. Nauta, "The Development, Evolution, and Status of Holland's Theory of Vocational Personalities: Reflections and Future Directions for Counseling Psychology," J. Couns. Psychol., 2010. <https://doi.org/10.1037/a0018213>
- [17] R. Du Toit and G. P. De Bruin, "The structural validity of Holland's R-I-A-S-E-C model of vocational personality types for young black South African men and women," J. Career Assess., 2002. <https://doi.org/10.1177/1069072702010001004>
- [18] D. Guglielmi, F. Fraccaroli, and M. L. Pombeni, "Les intérêts professionnels selon le modèle hexagonal de Holland," L'orientation Sc. Prof., 2004. <https://doi.org/10.4000/osp.700>
- [19] A. R. Spokane, E. J. Luchetta, and M. H. Richwine, "Holland's theory of personality in work environments," in Career choice and development, 2002.
- [20] A. R. Spokane and M. C. Cruza-Guet, "Holland's Theory of Vocational Personalities in Work Environments," in Career development and counseling: Putting theory and research to work, 2005. <https://doi.org/10.4135/9781412963978.n537>
- [21] Tetreau.B, "The rise of a psychology of professional interests," Careerology, vol. 10, no. 1, pp. 77–118, 2005.
- [22] V. P. & Bernaud.J, "The interests of the French People are they hexagonal?" 1. Elements for the validation of the model of the interests of Holland (RIASEC) in France," Quest. Orientat., vol. 49, no. 1, pp. 17–39, 1994.
- [23] "Holland's Codes," [https://en.wikipedia.org/wiki/Holland\\_Codes](https://en.wikipedia.org/wiki/Holland_Codes). [Online]. Available: [https://en.wikipedia.org/wiki/Holland\\_Codes](https://en.wikipedia.org/wiki/Holland_Codes).

- [24] "Dictionnaire septembre des métiers et professions ; suivi du guide Cléo, des clés pour s'orienter.," Septembre editeur S.E.N.C. pp. 256–257, 2005.
- [25] A. Lally et al., "Question analysis: How Watson reads a clue," IBM Journal of Research and Development. 2012.
- [26] J. Silva, L. Coheur, A. C. Mendes, and A. Wichert, "From symbolic to sub-symbolic information in question classification," *Artif. Intell. Rev.*, 2011. <https://doi.org/10.1007/s10462-010-9188-4>
- [27] T. Lei, Z. Shi, D. Liu, L. Yang, and F. Zhu, "A novel CNN-based method for question classification in intelligent question answering," in *ACAI 2018 Proceedings of the 2018 International Conference on Algorithms, Computing and Artificial Intelligence*, p. 54. <https://doi.org/10.1145/3302425.3302483>
- [28] Z. S. Harris, "Distributional Structure," *WORD*, 1954.
- [29] P. Kashyap and P. Kashyap, "Machine Learning Algorithms and Their Relationship with Modern Technologies," in *Machine Learning for Decision Makers*, 2017. [https://doi.org/10.1007/978-1-4842-2988-0\\_3](https://doi.org/10.1007/978-1-4842-2988-0_3)
- [30] Hourrane, Oumaima, and El Habib Benlahmar. "Survey of plagiarism detection approaches and big data techniques related to plagiarism candidate retrieval." *Proceedings of the 2nd International Conference on Big Data, Cloud and Applications*. ACM, 2017. <https://doi.org/10.1145/3090354.3090369>
- [31] Hourrane, Oumaima, et al. "Using Deep Learning Word Embeddings for Citations Similarity in Academic Papers." *International Conference on Big Data, Cloud and Applications*. Springer, Cham, 2018. [https://doi.org/10.1007/978-3-319-96292-4\\_15](https://doi.org/10.1007/978-3-319-96292-4_15)
- [32] Omar Zahour, El Habib Benlahmar, Ahmed Eddaoui, "E-orientation: entre prescription des théories et prise de décision" "Conference TIM'16, 2016.
- [33] Omar Zahour, El Habib Benlahmar, Ahmed Eddaoui, "E-Orientation: Between prescription of theories and decision-making" "Conference SITA'16, October 2016.
- [34] Omar Zahour, El Habib Benlahmar, Ahmed Eddaoui, "E-orientation: Vers une modélisation des facteurs d'orientation scolaire" "Conference TIM'18, 2018.
- [35] Team, Azureml, Dorard, Louis, Reid, Mark D, Martin, Francisco J, "AzureML: Anatomy of a machine learning service in" *JMLR: Workshop and Conference Proceedings 50*:1–13, 2016.
- [36] Omar Zahour, El Habib Benlahmar, Ahmed Eddaoui, Oumaima Hourrane "Automatic Classification of Academic and Vocational Guidance Questions using Multiclass Neural Network," *International Journal of Advanced Computer Science and Applications* 10(10);2019 <https://doi.org/10.14569/ijacsa.2019.0101072>
- [37] Zheng S., Yang M. (2019) A New Method of Improving BERT for Text Classification. In: Cui Z., Pan J., Zhang S., Xiao L., Yang J. (eds) *Intelligence Science and Big Data Engineering. Big Data and Machine Learning. IScIDE 2019. Lecture Notes in Computer Science*, vol 11936. Springer, Cham. [https://doi.org/10.1007/978-3-030-36204-1\\_37](https://doi.org/10.1007/978-3-030-36204-1_37)
- [38] Ekanayake, S. Y., & Samarakoon, K. (2016). Support of Mobile Phones in a Private Network for Science Teaching. *International Journal of Interactive Mobile Technologies*, 10(2), 4-9. <https://doi.org/10.3991/ijim.v10i2.4817>
- [39] Jeljeli, R., Alnaji, L., & Khazam, K. (2018). "A comparison between moodle, Facebook, and paper-based assessment tools: Students' perception of preference and effect on performance". *International Journal of Emerging Technologies in Learning*, 13(5), 86–98. <https://doi.org/10.3991/ijet.v13i05.8091>
- [40] Astalini Astalini, Darmaji Darmaji, Wawan Kurniawan, Khairul Anwar, Dwi Agus Kurniawan "Effectiveness of Using E-Module and E-Assessment". *International Journal of*

interactive Mobile Technologies (IJIM),vol 13 no 09 ,2019, <https://doi.org/10.3991/ijim.v13i09.11016>

- [41] Evelyne Kasongo Kkonko, Norman Chiliya, Tinashe Chuchu, Tinashe Ndoro “An Investigation into the Factors Influencing the Purchase Intentions of Smart Wearable Technology by Students “ International Journal of interactive Mobile Technologies (IJIM),vol 13 no 05 ,2019, <https://doi.org/10.3991/ijim.v13i05.10255>.

## 8 Authors

**Omar Zahour** holds a Master's degree in computer science of systems distributed at the University of Sciences in Agadir, University Ibn Zohr in Morocco; He is also a PhD student in Mathematics and Computer Science and Information Processing at Laboratory of Information Technology and Modeling, Hassan II University, Faculty of Sciences Ben M'SIK, Casablanca, Morocco. His research interest includes E-Oriented Systems, Semantic Web, Machine Learning, Data Science and Big Data. Email: [orzahour@gmail.com](mailto:orzahour@gmail.com)

**El Habib Benlahmar** holds his doctorate in computer science, Now he is a teacher of higher education at Ben M'Sik Faculty of Science, Laboratory of Information Technology and Modeling, Hassan II University, Casablanca, Morocco. El Habib Benlahmar has published several publications in various international journals and at national and international conferences. Among his research interests are:

Machine Learning; E-Learning; Cloud Computing; Data Science; Ontology; Deep Learning; Internet of Things; Semantic Web; Mathematics; Semantic Web Technologies; Mobile Applications; Educational Technology; Human-Computer interaction.

**Ahmed Edaoui** holds a doctorate in computer science, now he is a teacher of higher education at the University of Sciences Ben M'sik, Laboratory of Information Technology and Modeling, Hassan II University, Casablanca, Morocco.

**Ahmed Eddaoui** has published several publications in various international journals and in national and international conferences. Among her areas of research are:

Information and Communication Technology; Computer Networking; Network Security; Network Communication; Security; Networking; Information Security; Computer Networks Security; Network Security; Network Architecture; Network Security; Network Technology.

**Oumaima Hourrane** is a Ph.D. student in Mathematics and Computer Science and Information Processing at Laboratory of Information Technology and Modeling, Hassan II University, Faculty of Sciences Ben M'SIK, Casablanca, Morocco.

Article submitted 2020-01-03. Resubmitted 2020-02-14. Final acceptance 2020-02-19. Final version published as submitted by the authors.