

Usage of Visual Analytics to Support Immigration-Related, Personalised Language Training Scenarios

Gerasimos Antzoulatos¹[0000-0002-7424-8928], Athanasios Mavropoulos¹[0000-0002-7326-5910], Grigorios Tzionis¹[0000-0002-0456-8280], Anastasios Karakostas¹, Almudena Gonzalez Costas², Marta González Burgos², Stefanos Vrochidis¹[0000-0002-2505-9178], Ioannis Kompatsiaris¹[0000-0001-6447-9020]

¹ Centre For Research & Technology Hellas - CERTH, Thessaloniki, Greece

² Metodo Estudios Consultores, Vigo, Spain

Abstract. Most language learning applications are aimed at students or people who already know a language and want to improve their skills, or want to learn a new language. These applications, while seeking to be interactive, are not aimed at immigrants, refugees or asylum seekers, since the latter have different needs and interests from casual learners, opting for language skills that will allow them to function independently in the host society. This research is part of the European project WELCOME, which seeks to use state-of-the-art technologies, such as Virtual Reality (VR) apps and dialogue agents, to support the reception and integration of Third Country Nationals (TCNs) in Europe. The platform will be tested in three languages (Catalan, German and Greek), in real situations that the TCNs, mostly immigrants, refugees and asylum seekers, face, combining linguistic activities, where aspects related to language and culture are worked on. Furthermore, a Visual Analytics Component (VAC) leverages authority (NGOs/State institutions) users' perceptual and cognitive abilities by employing interactive visualisations as interfaces between users and learning analytics outcomes generated by amassed data. The goal is to find patterns within the characteristics of TCNs, and thus help language teachers adapt the content and tools to TCNs, contributing to greater personalisation in learning.

Keywords: Visual Learning analytics; Educational Data Mining; Language Learning; Language Teaching; Third Country Nationals; Migration

1 Introduction

In the last few decades, with the rise of IoT technologies, e-learning, online/mobile learning environments and tools have proliferated. Hence, new opportunities for supporting learning processes across all educational aspects have emerged that provide learning experiences (synchronous or asynchronous, remote or collaborative) in multiple participants, such as groups of students, individuals and teachers. This growth and widespread use results in the rapid increase of learning data that are gathered in educational institutions. Currently, the Educational Data Mining and Learning Analytics

fields encompass interdisciplinary methods to analyse and visualise data during teaching, learning, as well as education administration and services. Moreover, the utilisation of advanced visual analytics technologies in learning and teaching analytics is inevitable due to the heterogeneity, complexity, temporal and unstructured nature of learning data [1].

These applications, while seeking to be interactive, are not aimed at the public of immigrants, refugees or asylum seekers who have other needs and interests. Most language learning applications are aimed at students or people who already know a language and want to improve their knowledge and skills or want to learn an additional language after mastering the second one. However, TCNs arriving in Europe do not constitute an homogeneous group, since they have diverse educational, linguistic and socio-cultural backgrounds. This heterogeneity of profiles is transferred to the classrooms and in this context, the role of the teacher becomes more demanding making the adoption of advanced analytical tools mandatory to assist in the monitoring of students' performance. In addition, TCNs are a group that needs to quickly develop a range of language skills in order to be able to function autonomously in the host societies, so the teaching of these skills should help them cope with everyday life situations. Also, as studies have pointed out [2], certain language learning applications, VR-based ones in particular, offer an opportunity to combine real spaces and situations, apply Artificial Intelligence (AI) solutions and introduce cultural themes in specific settings, making the general context much closer to reality.

The aforementioned aspects constitute an innovation within the WELCOME project which aims at the development of a useful tool to provide, among others, the means to assist in the reception and integration of immigrants and refugees in host societies via educational content and learning (e.g. visit to the doctor, municipality application process, employment services, schooling system). Moreover, the cognitive and perceptual skills of authority users from NGOs and State institutions that offer language courses are enhanced via the support of the VAC, which empowers them to leverage interactive visualisations as interfaces between users and learning analytics outcomes generated by amassed data. By exploring possible correlations and patterns encountered within the characteristics of TCNs, language teachers may adapt the available content and language tools to each TCN separately, accomplishing the goal of personalised learning.

The main contribution of this work is, thus, twofold. Initially, a presentation of a contemporary language learning approach that aims to contribute to personalised training, customised to TCNs' needs and profiles is offered. It is then complemented by an introduction to the knowledge generation visual analytics framework which materialises in the VAC, conceived as a tool focused towards language teachers. Specifically, details are shared on how the scenarios and materials developed in the WELCOME project are designed to develop the oral skills of TCNs in the host language, with special emphasis on those situations and vocabulary that are most needed in the early reception phases; with emphasis on the oral interaction and simulated dialogues combined with activities for vocabulary acquisition. Furthermore, an actual immigration-related, language learning scenario is presented, along with the respective visualisations that the VAC offers, to better convey the usefulness of the overall approach.

2 Related Work

2.1 Language Learning

The increasing number of migrants in Europe have turned language classrooms into heterogeneous places, in which populations converge that differ in age, gender, socio-cultural background, first language, motivation and exposure to the target language. The role of teachers in these educational contexts is more complex and demanding than in regular, single ethnicity, language teaching contexts, and the teachers need to seek specialisation in working with migrants and refugees.

Adult foreign language learners have specific needs and characteristics that must be taken into account to develop materials that are tailored to these needs. Knowles [3] identified some characteristics of adult learning in his theory: adult learners are autonomous and self-directed, accumulators of life experience and knowledge, goal motivated, guided by relevance, practical and feel a greater need to be respected. Furthermore, foreign languages are sometimes taught in a way that is not satisfactory enough for migrants as the curricula includes themes that are not to their interest and training exercises do not refer to real life situations.

Scenario-based teaching for adult migrants. In 2012 Switzerland presented *fide*, the innovative conceptual framework for the linguistic integration of migrants in Switzerland. *Fide* stands for français, italiano, deutsch which are the languages that migrants have to learn in various parts of the country. Frequent contact situations between migrants and Swiss residents have been identified and analysed in different everyday situations, such as contact with authorities, work context, and health. These formed the basis of an inventory of “scenarios” which included: descriptions of interactive situations, the interlocutors involved, their respective roles, the overall aim to be achieved by the interaction, the usual course of action, socio-cultural factors to be considered, and helpful linguistic resources to achieve the interaction aim.

Task-based and action-oriented training programs for migrants. One of the pillars that sustains the Common European Framework of Reference (CEFR) is the action-oriented approach. CEFR suggests that the lessons and language training courses should be planned backwards, from the learner's real-life communicative needs, with a consequent alignment between curriculum, teaching and assessment [4].

From 2010 to 2015, the University of Thessaly implemented two nationwide training courses for newly arrived migrants, ELMEGO and MATHEME. The ELMEGO project was aimed at parents with children attending compulsory education, while MATHEME was aimed at unemployed migrants. In both cases the task-based learning methodology was adopted and an in-depth qualitative analysis related to immigrants' communicative needs (in everyday life and work activities) was conducted. In both cases, the task-based approach was proven to be very effective and furthermore resulted in team-building, identity investment, and empowerment.

2.2 Visual Analytics

Although our ability to gather and store massive data from heterogeneous sources has been reinforced, our ability to analyse and utilise them for making efficient decisions

is not so developed. Making matters worse, the analysis of such amounts of disparate data can lead to the well-known problem of "information overload" (also known as the "data deluge" problem) which concerns the danger of getting lost in data due to the irrelevant to the current task, the inappropriate (system-driven) processed way, or the illustration of the results ([5], [6]). Hence, to tackle this problem, novel visual analytics approaches engage human cognitive interactions and reasoning, along with advanced analytics processes. As Endert et al. [7] pointed out a shift from a "human in the loop" to a "human is the loop" viewpoint, which empowers visual analytics capabilities, with the human intuitive capabilities of interactive visualization [8].

Therefore, to discover hidden knowledge from massive, heterogeneous and complex data, firstly James Joseph Thomas established the Visual Analytics (VA) field, which was defined as "*the science of analytical reasoning facilitated by interactive visual interfaces*" [9]. Keim et al. (2008) defined VA as the process that incorporates advanced automated analysis techniques along with interactive visualizations aiming at effective understanding, reasoning and decision making on the basis of very large and complex data sets" ([5], [10]). In [6], Cui proposes a more detailed and comprehensive definition of VA in which the interactive visualizations, algorithmic data analysis and analytical reasoning techniques encapsulate human judgment into the KDD process to visually discover explainable patterns (knowledge) and to gain insight into large and complex data sets.

The aforementioned definitions for VA imply the solid interplay between visualisation, human intelligence, cognition and perception abilities, along with the advanced analytical processes performed by computers to obtain seamlessly, meaningful and explainable results leading to the generation of valuable knowledge for decision-making ([8], [11], [13]). Sacha et al. in [8] established a knowledge generation model for visual analytics which relied on the visual analytics process proposed by Keim et al. ([5], [6], [12]) and simultaneously it encapsulated human perceptual and cognitive theories, such as sensemaking [11] in order to generate knowledge in an iterative manner.

In [13] authors conducted a comparative review of state-of-the-art commercial VA systems. During the evaluation process, the VA tools were compared in order to cover the three main actions in a VA system workflow, namely data management, automatic analysis, visualisation and system and performance. It is worth mentioning that the majority of the visual analytics systems do not fully cover or appropriately deal with the aspects of the Knowledge Generation model [8].

3 Proposed Approach

3.1 Motivation for the Creation of Scenarios

The creation of language learning scenarios has been approached from a co-creation methodology viewpoint with the teachers and Public Administration users, in combination with an in-depth study of the applications and resources on the market for teaching languages to migrants, adults and applications that make use of virtual reality and visual analytics for language learning. Moreover, since the teacher's role is fundamen-

tal to adult learners with an A1-A2 level, all scenarios have been designed to complement face-to-face language lessons and to allow TCNs to practice the most likely situations they will encounter during the different stages from arrival to integration.

Several apps for language learning have been tested, including the most popular on the market: Duolingo, Drops, Word of the day (iOS versions), Babbel, and Mondly (web versions). These are focused on autonomous learning and require some previous knowledge on the target language. The main problems detected are the following:

- Lack of spelling activities and/or lack of activities to learn the alphabet. Of all the apps analysed, only Drops has an activity aimed at learning the alphabet. While trying to learn a language with an alphabet different from the native one, a well-planned activity for learning the alphabet becomes a crucial activity for being able to advance to more complex lessons.
- Limit of lessons or errors per day: users in the free version of Duolingo, Babbel, Drops and Mondly have a limited number of lessons or errors in the activities each day. When users reach that limit, the app is blocked until the next day. As beginners make mistakes frequently, if the application does not allow them to finalise a lesson, it can demotivate them and they may not access the course again.
- Free navigation: Duolingo, Mondly and Babbel do not allow participants to freely navigate the courses. They have to follow the training path, as the free versions don't offer any possibility to personalise the training path. If students come across several topics in a row that are not to their interest, they can drop out of the course.

3.2 Building the Learning Program: Scenario-based teaching

The methodology chosen for language learning is the scenario-based teaching for its proven efficiency, which in addition offers the best development options in VR. The teachers who will pilot the program have also expressed the usefulness that their students find in the programs based on this approach.

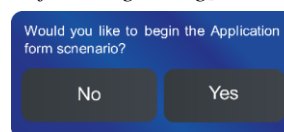
Sheridan and Kelly [14] claim that scenarios should be related to the real world, so learners could find a connection between contents and application of such contents in their lives. In order to achieve the objective of enabling TCNs to cope with everyday situations, language lessons are developed, linked to respective situations in which cultural aspects specific to each country are introduced. Thus, in a self-presentation scenario, the person is expected to be able to provide basic information about himself/herself, but also to understand the cultural differences in aspects such as surname or family name depending on whether one is single or married. Hence, each language learning scenario is composed of several activities: a simulated dialogue between the TCN and a 3D avatar to train specific vocabulary, easy grammatical constructions and/or cultural lessons associated with that scenario.

The scenario to be tested refers to the “First Reception Service” (FRS) and the main aim is to train TCNs in providing basic information about themselves. The human-machine interaction is handled by the dialogue agent which requests the required information from the TCN and then allows the latter to test the procedure and acquired knowledge via VR minigames. The vocabulary and grammatical constructions of this scenario correspond to levels A1-A2 of the CEFR. It is practiced with the help of the

teacher, who will assist TCNs to log into the WELCOME platform and put on the VR headset. Once the TCNs access the VR platform, users are placed in an apartment-shaped stage, in which they are greeted by a personalised avatar that explains the instructions to access the learning scenarios.

The personalised avatar will provide each TCN the following instructions to complete the FRS scenario: “You need to complete an application form. Take a look around your apartment for the application form, once you have located it, pick it up using your controller to enter your first learning simulation” [*application form is glowing*].

Once the TCN has located the application form, a floating UI appears to ask the TCN to start the language-learning scenario. When the TCN presses Yes, an office scene is loaded, where the user has to use the controllers to navigate the VR office and enter in the VR rug. Once the TCN enters the VR rug, the personalised avatar will greet her/him, the questions will pop-up as a floating UI and the TCN will be asked to choose the correct answer for each of the questions posed by the avatar. The first floating UI will appear in their native language and it will read: “Please provide an appropriate response according to the correct time of day by pressing one of the options below”.



For each of the questions that are part of each scenario, the system follows the same flow: once TCNs obtain the score of the question, the avatar will move to the next question. TCNs have 3 attempts to answer the multiple-choice questions.

Level 1. TCNs are able to provide a correct response without any explanation. First attempt. Score 10

Level 2. TCNs are able to provide a correct response with an explanation about the meaning of the word. Second attempt: Score: 5

Level 3. TCNs need the word in their mother tongue. Third attempt: Score 0, if the TCNs provide a wrong response, and score 1, if the TCNs provide the correct response.

The total mark is between 0 and 10 for each question of the dialogue and the final mark will be the average of all the dialogue parts. The class mark is the average of the different exercises (in total, and by activity), so the teacher has information to provide recommendations that will reinforce contents not yet fully assimilated by the TCN.

3.3 Methodological approach for testing language learning scenarios

The vocabulary and grammatical constructions of each scenario are grouped by CEFR levels. For the first pilots with TCNs, the language learning materials only target A1 and A2 levels, but the levels will be expanded in future scenarios to achieve a B2 or C1 level. To collect crucial users' data in order to establish correlation and patterns, socio-demographic information (age, gender, country of origin, mother tongue, educational level) and technological skills or knowledge of other languages will be requested while creating their profile in the WELCOME platform. The VAC will be in place to tackle the analysis of the aggregated information originating from the language learning scenarios and user profile, and provide the opportunity to the interested party, be it a teacher or an authority user, to draw important conclusions that will assist in forming a

personalised educational strategy. The suggested solution relies on the adoption of Visual Data Analytics techniques that enable the discovery and understanding of patterns, advanced correlations and trends in large datasets, generated by learning scenarios, via visual interpretations. In this stage, users will be able to utilise advanced techniques, integrated into the VAC and the following correlations are expected to be possible:

- Valuation differences in relation to gender, age, educational level, previous digital skills, and languages known to immigrants, refugees or asylum seekers;
- Differences in the results obtained in the learning activities in relation to the aforementioned attributes;
- Differences in the learning pace of the TCNs participating in the pilots, depending on where the pilots will take place, Spain, Greece and Germany;
- Differences in performance between the TCN and their own classmates;
- Differences in performance between the TCN and other TCNs with a similar profile, including % of TCNs unable to finalise the training activities and VR scenarios.

Given the differences in the profile of the TCNs that are going to participate in the pilots, respective variations are expected in the learning pace and in the scores obtained in the different scenarios. TCNs who have been exposed to the language of the host country or whose native language shares the same alphabet with the language they are learning, will advance more quickly and will obtain higher scores in A1-A2 scenarios than those who do not know the language or who have to learn the alphabet.

It should be mentioned that great care has been taken to prevent unauthorised access to private and sensitive information concerning the TCNs by providing a secure environment, protected by specific access rights/privileges, based on the user's role. Thus, unauthorised people cannot access the TCNs' personal data or their language learning performance. Therefore, only teachers can carry out relative comparison analyses and, specifically, only concerning the performance of TCNs of their own class.

3.4 Visual Analytics for Language Learning Scenarios

In this section, we will propose a Visual Analytics schema that enables practitioners to analyse information and foster complex decision-making processes in the language learning domain. It is motivated by the knowledge generation model (KGM) for visual analytics which has been proposed by Sacha et al. [8]. Similar to KGM, our proposed model consists of the main components that are Data, Data Mining Models, Visualisation and Knowledge. However, it encapsulates the Decisions component due to the fact that, in the framework of the language learning scenarios, the interest focuses on how the teachers will be assisted by it, in order to make efficient and timely decisions in terms of the TCN learning/educational process (see Fig. 1).

In this model, the knowledge can be generated by Visual Data Exploration going from data via visualisation (InfoVis pipeline) to knowledge and Automated Data Analysis that is going from data via models to knowledge (see Fig. 1).

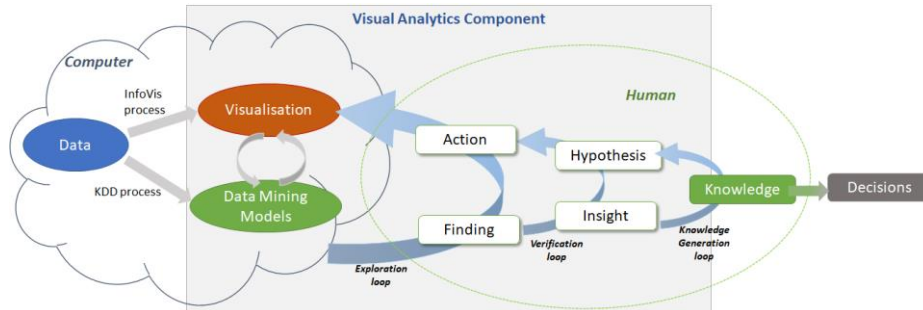


Fig. 1. Knowledge generation model in VAC for language learning

The computer part of this Knowledge generation model for Visual Analytics breaks down into Data acquisition processes, and pipelines from Data to Visualisation (InfoVis process) and Data to Model in terms of Knowledge Discovery in Databases (KDD) processes, coupling with machine learning algorithms. Data quality plays an important role in visual analytics processes and depends on how the data were generated, gathered, and selected for further analysis. During an analysis, aggregated data could be produced either by the application of automated methods (e.g. classification or clustering) or by manual annotations. The information visualisation pipeline employs techniques, usually from Exploratory Data Analysis (EDA), to detect relationships in the data in an illustrative manner. The main components of the InfoVis pipeline includes the transformations/mappings between Raw Data, Data Tables, Visual Structures and Views which can be manipulated through Human Interactions [8]. Furthermore, visualisations can be produced by automated models; for instance, data can be illustrated as groups after undergoing a clustering analysis. In particular, models can originate from descriptive statistical approaches or even more complex ones, coming from the KDD and Data Mining field. Complex patterns, relationships, and associations can be revealed and visualised, in order to be communicated and comprehended by the analysts. The KDD process consists of iterative and interactive steps, in which the data are selected and filtered, preprocessed, and transformed appropriately, in order to apply data mining techniques on them. The goal is to discover unknown patterns that are hidden in the data and convert them into valuable knowledge [8].

The exploration loop concerns the interaction between analysts and visual analytics systems to analyse data, generate new visualisations or models. Actions concern analysts' goals and tasks that produce tangible and unique responses from the visual analytics system. Actions derived from Hypotheses are usually complex Actions, while those that are derived from Findings are normally simple Actions, such as changing the mapping of visualisation or selecting a different feature for model building. Also, Actions can deal with data gathering or selection in order to prepare the data, or Actions to create models (model building) or to utilise existing models. Similarly, visual mapping Actions concern the creation of data visualisations, while model-vis mapping Actions map models into visualisations. Finally, Actions that enable analysts to manipulate with viewpoints focusing on interesting data in the visualisation and interact with the visual analytics system, are also provided [8]. This interaction can lead to emerging new interesting observations (Finding) or to new Insights. Missing or extreme values

can be considered as Findings in the data that affect the further analysis and hence require special data processing. In the case of visualisations or models, a Finding can be a pattern, a trend, an evident model result, even an unusual behaviour of the system. Findings can be derived either by the utilisation of automatic data mining models or by humans-analysts exploiting their visual perceptions and cognition skills. In general, the visual analytics systems rely on the Actions and Findings in the exploration loop, as the understanding and efficient interpretation of the Findings in the context of the problem domain, provide new Insights to analysts. Going a step beyond, Insights may lead to new Hypotheses which require further investigation to verify or falsify them, in the verification loop. This process, namely the assessment of the Insights, results in the gain of new Knowledge.

The visual analytics process facilitates analysts to identify evidence for existing assumptions or learn novel and valuable knowledge, so as to foster and enrich their prior knowledge related to the problem domain. The derived knowledge from visualisations, automatic analysis along the preceding interactions between visualisations, models, and the human, can be considered as *Prior Knowledge* [15]. The involvement of experts in the exploration of knowledge is significant as they will be immediately aware of the relationship between new knowledge and the existing domain knowledge. Complex reasoning and sense-making processes are employed as the basis for generating additional knowledge or can be composed with Prior Knowledge to produce more general truths enhancing the *User Knowledge* ([11], [16]). In the proposed model, the iterative process that addresses high level reasoning, such as inductive, deductive, and abductive reasoning, in the knowledge generation, exploration and verification loops, enable end-users to make Hypotheses and apply criteria to make decisions.

In the framework of the WELCOME project, the VAC will be further developed serving the needs for language teachers, professionals from NGO's and public entities who are involved in the first reception and integration of TCNs in the host societies. Specifically, the VAC encompasses processes for data preparation, as well as information visualisation processes in an interactive, intuitive and user-friendly VAC UI aiming to engage humans in the loop of exploring data and the discovery of new knowledge. The VAC UI will be enriched with functionalities that enable authorities to visualise aggregated information related to health data, personal, educational and professional profile of TCNs. Furthermore, it will be able to synthesise information and derive insights from massive, dynamic and ambiguous data concerning TCNs and visualise normal and abnormal detected trends and patterns related with them.

Furthermore, more complex information generated by the high-level knowledge interpretation will be achieved by applying KDD processes and data mining. The outcomes of the analysis will be visualised and will be incorporated to the VAC UI aiming to enable end-users to improve evidence-based decision making. Hence, the exploration loop is supported by providing customised visualisations for different research analytical questions raised by end-users. Moreover, VAC tailors functionalities that allow end-users to tune parameters, choose specific characteristics and in general take actions that naturally provoke interactions of analysts with the system. VAC provides the tools to transforming the Findings from the exploration loop to Insights by verifying or falsifying concrete Hypotheses. This knowledge generation process will be capable of

achieving sustainable decisions for migrant management and integration as well as for the development of their language learning skills.

3.5 Evaluation Visual Analytics process

As mentioned above, the utilisation of VAC in language learning can reveal valuable knowledge concerning the progress of TCNs or to detect the difficulties that TCNs may encounter in following or completing the language teaching scenarios. Hence, a teacher, via the VAC UI, will be able to visualise correlations among students' profiles that are participating in a specific class and follow the activities of the particular language learning scenario. In Fig. 2 an example dashboard created by a teacher of a class (C2) is exhibited. Using visual analytics tools, he/she can graphically present combinations of characteristics and socio-demographic information among students in C2.



Fig. 2. Dashboard illustrating correlations among general characteristics of class C2 students

Furthermore, the VAC enables the teacher to compare the student's performance with the average performance of his/her class, as illustrated in Fig. 3, which includes:

- General TCN's personal information, as well as some basic statistical measurements in terms of his/her performance. The VAC provides statistical data as MAX, MIN, Median, Mean, Standard deviation in terms of the score achieved by specific TCNs in the activities.
- Average score of the students depending on the different scenario activities and compared to the specific TCN (*Average Performance* bar plot).
- Deviation between the particular student and the average score of the class per activity (radar plot).
- % of students that are able to finalise the activities in the TCN's class (pie chart).
- Three barplots which exhibit the average performance of TCNs in a specific class over activities grouped by gender, time needed to complete an activity discretised in days and educational level of TCNs.

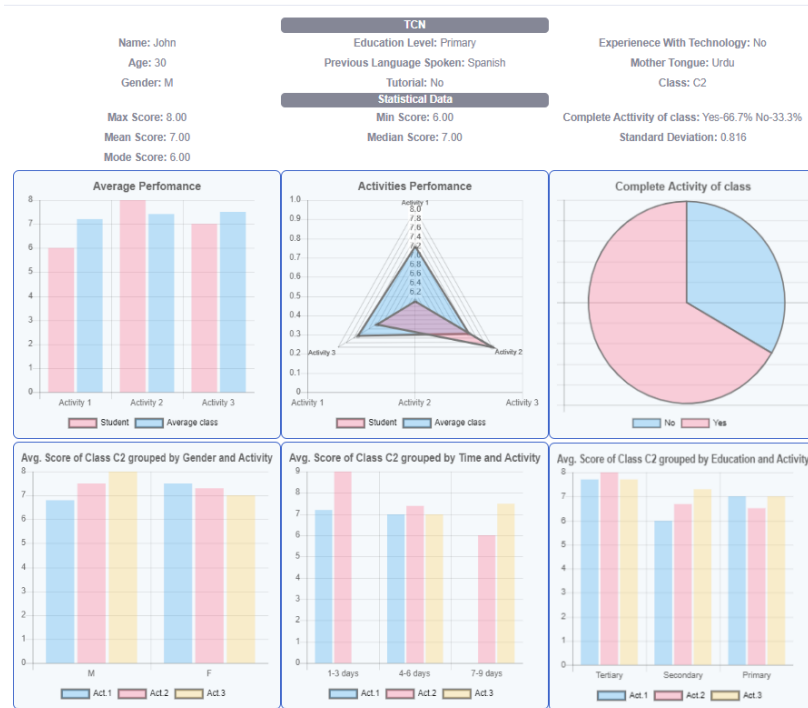


Fig. 3. Performance evaluation of TCN comparing with the average class performance

4 Conclusions and Future Work

In this work, we propose an approach that allows teachers to concisely evaluate the evolution of students in complex interaction environments through advanced metrics that go beyond the usual metrics of learning platforms such as connection time or the results obtained in the evaluation tests. Simultaneously, a visual analytics knowledge generation framework is proposed, which encapsulates data analytics techniques and sophisticated visualizations that enable teachers to gain insights concerning TCNs' language learning performances. The ultimate goal is to empower teachers and authorities to detect problems that TCNs can encounter, and anticipate them in time to effectively support the authoring of appropriate language courses.

As future directions, the proposed framework will be soon evaluated over real use cases and should be further developed according to the provided feedback. Additionally, reinforcing the VAC with enhanced data mining processes, such as classification/clustering algorithms and association rules, will provide a robust environment to reveal hidden and valuable knowledge to improve the TCNs' language learning process. For example, creating profiles or groups of TCNs in the same classroom with similar characteristics in terms of their activity scores and the automated correlation with socio-demographic information among those students, will allow the system to provide to the teachers personalised, student-specific recommendations.

Acknowledgements

This work has received funding by the EC under contract H2020-870930 WELCOME.

References

1. Vieira, C., Parsons, P., Byrd, V.: Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education* 122, 119-135 (2018).
2. Reyes Sánchez, B. and S.: A Review of Virtual Reality-Based Language Learning Apps. *RIED. Revista Iberoamericana de Educación a Distancia* 24(1), 159-177 (2021)
3. Blatyta, D. F., Spada, N. M.: O papel do professor de línguas na construção de uma aprendizagem significativa. In: Silva, Kleber Aparecido; Alvarez, Maria Luisa Ortiz (Eds.), *Perspectivas de investigação em Linguística Aplicada*. São Paulo: Pontes (2006)
4. Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. Common European Framework of Reference for Languages: learning, teaching, assessment. Cambridge University Press (2001).
5. Keim D., Andrienko G., Fekete JD., Görg C., Kohlhammer J., Melançon G.: Visual Analytics: Definition, Process, and Challenges. In: Kerren A., Stasko J.T., Fekete JD., North C. (eds) *Information Visualization. Lecture Notes in Computer Science*, vol 4950. Springer, Berlin, Heidelberg (2008). DOI: https://doi.org/10.1007/978-3-540-70956-5_7
6. W. Cui: Visual Analytics: A Comprehensive Overview. In *IEEE Access* 7, pp. 81555-81573 (2019). DOI: 10.1109/ACCESS.2019.2923736.
7. Endert, A., Hossain, M., Ramakrishnan, N., North, C., Fiaux, P., Andrews, C.: The human is the loop: new directions for visual analytics. *Journal of Intelligent Information Systems*, pp. 1–25 (2014).
8. Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis G., Keim, D.A.: Knowledge Generation Model for Visual Analytics. In *IEEE Transactions on Visualization and Computer Graphics*, 20(12), pp. 1604-1613 (2014). DOI: 10.1109/TVCG.2014.2346481.
9. Thomas, J.J., Cook, K.A.: *Illuminating the Path*. IEEE Press (2005).
10. Thomas, J.J., Cook, K.A.: A visual analytics agenda. *IEEE Computer Graphics and Applications* 26, pp. 10-13 (2006).
11. Pohl, M., Smuc M., Mayr, E.: The User Puzzle—Explaining the Interaction with Visual Analytics Systems. In *IEEE Transactions on Visualization and Computer Graphics* 18(12), pp. 2908-2916 (2012). DOI: 10.1109/TVCG.2012.273.
12. Keim, D., Kohlhammer, J., Geoffrey, E., Florian, M.: *Mastering The Information Age – Solving Problems with Visual Analytics*. Eurographics Association (2010).
13. Zhang, L., Stoffel, A., Behrisch, M.: Visual analytics for the big data era — A comparative re-view of state-of-the-art commercial systems. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 173-182 (2012). DOI: 10.1109/VAST.2012.6400554.
14. Sheridan, K. M., Kelly, M. A.: Teaching early childhood education students through interactive scenario-based course design. *Journal of Early Childhood Teacher Education*, 33(1), 73-84 (2012).
15. Endert, A., Ribarsky, W., Turkay, C., Wong, B.L.W., Nabney, I.T., Diaz-Blanco, I., Rossi, F.: The State of the Art in Integrating Machine Learning into Visual Analytics. *Computer Graphics Forum*, 36(8), pp. 458–486 (2017). DOI:10.1111/cgf.13092.
16. Pirolli, P., Card. S.: The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis*, volume 5, pp. 2–4 (2005).