Pavithra.K, International Journal of Computer Science and Mobile Applications,
Vol.5 Issue. 11, November- 2017, pg. 95-102 **ISSN: 2321-8363**
**Impact Factor: 5.515**

# A Study on Supporting Visual Narratives Student Engagement using Big Data Technologies

**Pavithra.K**, Msc (PhD),
Assistant Professor, KG College of Arts and Science, India
kpavithramsc@gmail.com

***ABSTRACT:** A novel visual narrative framework that has been exposed to facilitate, support and enhance student commitment in an adaptive Online Learning Environment (OLE). VisEN provides explorable visual narratives modified to students in order to support them to engage with course content. The evaluation of VisEN showed that the explorable visual narratives confident the majority of improving engagement students that completed the Information Management and Data Engineering module as part of their undergraduate degree, to engage with assigned activities , and subsequently these learners enhanced their engagement levels. It might make the power system load varied complex than before which will bring difficulties in short-term load forecasting area. To overcome this issue, this paper proposes a new short-term load forecasting framework based on big data technologies. First, a cluster analysis is performed to classify daily load patterns for individual loads using smart meter data. Next, an association analysis is used to determine critical influential factors. This is followed by the application of a decision tree to establish classification rules. Then, appropriate forecasting models are chosen for different load patterns. Finally, the forecasted total system load is obtained through an aggregation of an individual load's forecasting results. Case studies using real load data show that the proposed new framework can guarantee the accuracy of short-term load forecasting within required limits.*

***Keywords*— *Association analysis, cluster analysis, short-term load forecasting. Personalized E-learning, Information visualization***

## INTRODUCTION

As Online Learning Environments (OLEs) grow in popularity student engagement with such environments remains an open issue. The literature has highlighted that engagement decreases over time in OLEs and overall is lower when compared to traditional classroom learning Student engagement is a key factor for development and learning hence this research focuses on supporting learner engagement when using OLE s This section introduces explorable visual storytelling to support learner engagement. Information Visualization (IV) facilitates an effective means to comprehend data by supporting pattern discoveries and the communication of data Storytelling in IV or visual narratives can be defined as an ordered sequence of steps consisting of visualizations, which are linked to make the communicated message more memorable. Another issue is that the load needs to be forecasted at the substation or bus level for calculation of the power flow. Most utilities do not process load forecasting at the substation or bus level because of the complexities involved in capturing the necessary information or because there is very little data available.

The construction of the smart gird has led to wide deployment of smart meters [18]. Smart meters provide an opportunity to analyse the time-varying characteristics of individual loads.

This research aims to communicate a personalized story to students by adopting a slide-show type approach (typically used in visual narratives in online journalism). Each slice or tab consists of a single visualization with a description to communicate part of the narrative. Collectively all the narrative slices communicate the entire story.

This is followed by the evaluation of the framework, which focuses on the impact the visual narratives had on student course engagement, learner perceptions, perceived understandings of the message, and narrative usage patterns. VisEN was used during two successive academic years in Trinity College Dublin to provide visual narratives to a total of 233 undergraduate students studying the Information Management and Data Engineering module. This module consisted of two parts: one part consisted of lectures on relational database management systems delivered by the module professor (not the focus of this research). The other part involved students studying Database development and SQL programming using the AMAS adaptive OLE. Apart from being introduced to AMAS and given a demonstration of it, the students were not provided with any other form of sup-port when working through their activities using the OLE. VisEN was deployed to AMAS and provided students.

## I. RELATED WORK

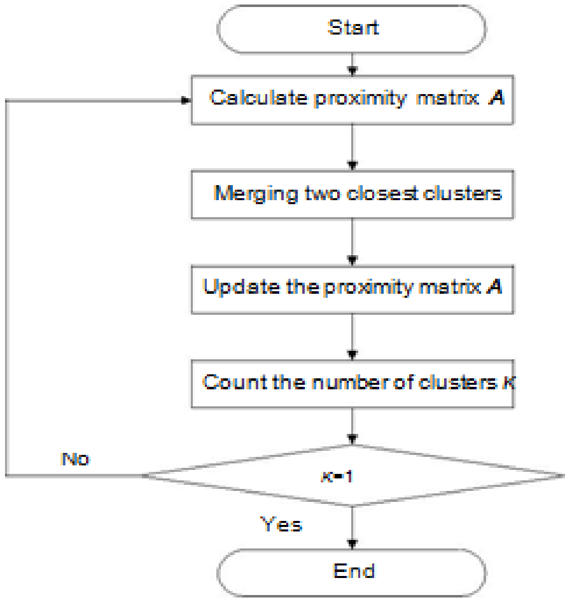
Visualizations are commonly used in OLEs within the Learning Analytics (LA), Open Learner Modeling (OLM) and Educational Data Mining (EDM) domains. In OLM, visualizations are used to present student models showing learner competencies to support reflection. EDM uses visualizations to present patterns of sequences to enable pre-dictions to be made. LA uses visualizations to present student activity to educators and at times to students to raise awareness and to promote behavioral change. This section analyzes the state of the art in OLEs, specifically in do-mains that present data to learners using visualizations (LA, OLM, and EDM) by examining the level of guidance and visual scrutability available to students.

## II. FORECASTING FRAMEWORK AND CORE TECHNIQUES

The proposed forecasting framework includes five steps, Steps 1 through 3 consist of machine learning techniques that aim to discover typical load patterns of historical load data and then find their critical influential factors to establish classification rules. Step 4 is a model training process, where parameter combinations for corresponding load patterns are chosen to build forecasting models. In Step 5, individual load forecasting consequences are added to appear at the final system load.

### A. **Identification of Critical Influencing Factors**

Many factors, including temperature, humidity, day type, etc. have impact on loads. For different loads, critical influential factors are not the same. For example, temperature could be a major influential factor for residential loads. But temperature may have little impact on some temperature non-sensitive industrial loads. Grey correlation analysis is functional to determine the critical important factors of each individual load.




?. Flow chart of hierarchical clustering algorithm.

In recent years, an increasing interest in visual narratives in IV and online journalism has emerged to support data analysis. In IV, tools such as Ellipsis, Tab-leau, SketchStory and Gapminder have successfully presented data through visual narratives and their evaluations have shown encouraging results.

To date, visual narratives have not been used to present learner data to students or educators. VisEN enables narrative authors to construct a story, which is automatically visualized, personalized and made explorable. The visual narrative consists of a personal message, displayed across multiple screens that can have a beginning, middle and end in order to guide the learner. Each narrative slice of the story focuses on guiding the student through the data presented by the visualization via a textual description.

## B. **Progressing the State of the art**

This research takes the opportunity to introduce visual narratives to OLEs, specifically adaptive OLEs. The visual narratives support interactions to assist students in analysing the message and include slice transformations to enable them to explore data directly related to the individual narrative slices.

The data transformation approach implemented by VisEN is different from those found in the LA, OLM and EDM domains as each visualization in the visual narrative can be explored through multiple transformations of the data used in the narrative slices. This approach enables students to explore each narrative slice of the visual narrative which can be useful to understand the message communicated. Slice transformations are more appropriate for presentations that display a flow of individual visualizations, one after the other (visual narratives) where each view can be explored as opposed to dashboards, where one or several visualizations are available on a single screen.

### C. **Narrative Consumption**

The Visual Narrative Explorer component enables students to view, interact and explore their visual narratives. This component personalized the visual narratives for each student by 1) updating the data used by the course instructor in the visual narrative to use current student logged data, and 2) personalizing the descriptions in the message. Fig. 3 presents screenshots of a part of a visual narrative of a student enrolled in the course during the 2014-2015 academic years. The students' visual narratives consisted of three narrative slices: 1) Completed Tasks, which enabled learners to examine their tasks and the times they spent on them; 2) Engagement Breakdown, which allowed students to analyse their engagement per task (Fig. 3A); and 3) Re-source Usage presenting the material they used and shared. Combined, these slices presented a personalized story with personalized descriptions to each learner. Each narrative slice consisted of two-three slice transformations (links within the slice), which when selected displayed popup windows enabling students to scrutinize and explore the story through related data

### D. **Research approach**

The aim of this research was to present personalized exportable visual narratives to students and then to evaluate 1) the impact (if any) that these visual narratives had on student course engagement, 2) learners' perceptions to-wards them, 3) whether students were able to understand them, and 4) their usage patterns.

A detailed study was conducted during and after the course (across the two academic years) consisting of three sets of analyses as part of the evaluation.

This section discusses the research approach used in the study consisting of quantitative and qualitative analyses.

The analyses comprised of the following:

1. The collection of data included student interactions with their visual narratives, learners' responses to a post-course questionnaire, and their opinions to-wards their visual narratives. Data collection procedures were consistent across both academic years of the course. The course, visual narratives and the questionnaire were also unchanged. Data from all students who participated in the course across both years was collected and used in the analyses. In addition, the material from the course remained un-changed, including activities, books, and support.
2. The quantitative method examined the impact that the visual narrative usage had on course engagement through statistical measures. This was con-ducted by examining the student logged data, which consisted of almost 120,000 student interactions for both academic years. Over 10% of these interactions were visual narrative interactions. The quantitative data also examined students' responses to the post-course questionnaire statements. It also included the running of  tests for independence on the responses.
3. The self-reported usage data used by the study was verified using the student logged data as they were linked. Due to the nature of the course (students working on activities at best suited times) and the lack of access to the students following the completion of the course, it was not possible to conduct further forms of qualitative analysis.

a) data storage, data management, computation, and application server), which reads and analyses the data and forms them into a key/value.

b) Shuffle: Worker nodes (computer servers) redistribute data based on the output keys (produced by the "map()" function), such that all data belonging to one key is located on the same worker node. In this process, load data of different collection times with corresponding users will be stored in the same group.

c) Reduce: Worker nodes (computers or servers) now process each group of output data per key, and in parallel. This step aims at merging similar load data and eliminating data error.

## E. The technical architecture contains four layers

### A. Data Storage Layer

The data storage layer applies two technologies: Hadoop distributed file system (HDFS) and relationship database management system (RDBMS). HDFS is applied to store smart meter data and weather data. RDBMS is used to store consumer data and system topology.

### B. Data Management Layer

The data management layer applies Sqoop technology to convert data from weather, smart meter, and consumer usage into a Hadoop database (HBase). In HBase, data from each consumer's load is stored in a single table. HBase applies a key/value technique for quick and efficient access to target data. The row key includes user ID and date. The column key is time. The value consists of measured electricity usage, temperature, humidity, wind velocity, and day type.

### C. Computation Layer

MapReduce in Hadoop processes parallelizable problems across huge datasets by using a large number of computers (nodes). The process includes three steps: map, shuffle, and reduce, as shown as.

For data importing processes, corresponding steps are shown below:

Map: Historical data sets, including load data and weather data are stored in HDFS in plain text style. Then, a map function is carried out by each worker node .

### D. Application Layer

The application layer includes all the technologies described in Section II, i.e., cluster analysis, association analysis, decision tree, and forecasting methods.

## III. CONCLUSION

This paper proposes a new framework for performing short-term load forecasting using smart meter data based on big data technologies. Our proposed solution includes five steps:

1. Load identification of load patterns using clustering techniques,

2. determination of critical influencing factors using association analysis,

3. establishment of classification criteria using decision tree,

4. choosing appropriate load forecasting model for each load pattern and associated critical factors,

5. forecasting each individual load and calculating the OLE. The analysis of related work highlighted that visual narratives have never been used in OLEs and discussed how they could guide students through an intended message, which could be explored and scrutinized.

# REFERENCES

[1] B. Ginon, C. Boscolo, M. D. Johnson, and S. Bull. "Persuading an Open Learner Model in the Context of a University Course: An Exploratory Study",In International Conference on Intelligent Tutoring Systems, pp. 307-313. 2016.
[2] J M. Peter, S. Jan, and Z. Iveta, "Concept definition for big data architecture in the education system," in Proceedings of IEEE 12th International Symposium Conference, Herl'any, Slovakia, 2014, pp. 331–334. S.
[3] J. Grann, and D, Bushway, "Competency map: Visualizing stu-dent learning to promote student success." In Proceedings ofthe fourth international conference on learning analytics and knowledge, pp. 168-172. ACM, 2014.
[4] A. Deihimi, O. Orang, and H. Showkati, "Short-term electric load and temperature forecasting using wavelet echo state networks with neural reconstruction," Energy, vol. 57, no. 4, pp. 382–401, Aug. 2013.
[5] J. Guerra, R. Hosseini, S. Somyurek and P. Brusilovsky. An intel Essa, and H. Ayad, "Improving student success using predic-tive models and data visualisations." Research in Learning Technology vol. 20, 2012.K. E. Arnold and M. D. Pistilli, "Course signals at Purdue," Proceedings of the 2nd International Conference on Learning Ana-lytics and Knowledge,pp. 267-270. ACM, 2012.
[6] Nithya.A, "Internet Information System Based on Network Security Design". ISBN:978-93-80506-08-1., National Level Conference on Innovative in Information Technology

# Biography

**Pavithra K**, MSc (PhD), Assistant Professor in Computer Science at KG College of Arts and Science. She is pursuing Doctorate in GRD College of Science Coimbatore. She is very much interesting in Research in Data Mining and computer networking also in paper publication and attending Conference and Seminars.