# GREEN LIGHT: A CONTINUING PROFESSIONAL DEVELOPMENT WEB AND MOBILE APPLICATION IMPLEMENTING LATENT DIRICHLET ALLOCATION (LDA) ALGORITHM

Cindy April R. Dakis, MSIT - dakisleochico@gmail.com
Rosicar E. Escober, Ph.D. - rose_escober@yahoo.com
Polytechnic University of the Philippines - Open University System
Polytechnic University of the Philippines
Sta. Mesa, Manila, 1016, Philippines

## ABSRACT

Dissemination of information plays a vital role in order to uplift the lives of every Filipino, this entails the adaptation of the new information and to be cognizant to the changing world of the industry. Being at par with the industry standards and excellence shows the worthiness to every professional, the continuing strives for excellence of Filipino experts to become globally competitive never stops. Consequently, the fast-changing industry needs makes individual tends to seek a higher leaning in terms of technological and skills development and/or enhancements through educating themselves to different facets of trainings, seminars, conferences and workshops and for them to be authenticated by the PRC board. The question in mind is, what channel to be use? The authenticity of the channel being used? This is where usually the problem arises.

The Green Light Mobile Application determines the level of agreements to professionals in terms of continuing professional development thru CPD act of 2016 and the status of the CPD Credit Units. The services will be carried out by the implementation of the Green Light application for ease of use and convenience through a real time – interactive monitoring of CPD Units and a proper allocation of trainings, seminars, conferences and workshops to individuals who desires to renew their respective licenses. The information from CPD Provider and from the internet will be sent to their mobile accounts and other transactions like monitoring, knowing and notification will be accessible via mobile technology which gives individuals an abrupt information perceptive.

Professionals greatly needed an application to ease their access to the PRC services and seminars, socially, they need something that will provide them an easy way of accessing it, removing the barriers of issues and concern of affordability and accessibility.

**KEYWORDS** - Latent Dirichlet Allocation Algorithm, Continuing Professional Development, Big Data, Meta Data Search

## INTRODUCTION

As the innovation of technology is consistent, the people in the society adjust, relying more on technology, because it currently gives the ease of access and portability to almost everything, and it is possible through the use of mobile phones and the internet. In line with this, the researcher conceptualized an idea that will help the Filipino professionals to relief their concerns in terms of accessing the PRC's services, and the best possible aid for their problems and concerns is to create a mobile application that implements a certain algorithm.

The purpose of this study is to help the Filipino professionals to be more involved in attaining and doing continuous professional development, not only for their own benefit, but also for the Philippines labor power and economy. By analyzing the theoretical frameworks, reviewing other related studies and literatures, using questionnaires, formulating conclusions and recommendations, and by developing the application system tool, this study will not only help the Filipino professionals to be more updated and observant of their CPD status, but also to those governing and other related bodies to use this as their future reference for their own future planning considerations.

## LITERATURE REVIEW

### Continuing Education as a Whole

Studying never ends after graduation. Meaning, those people who are professionals or to those who are working in the field of their expertise never stops studying, learning or adapting new skills and knowledge relating to their profession. Others may see it as a waste of time, or unnecessary, but continuing education as a professional is important, particularly to those professionals who needs to adapt to the new trends of their profession, and also to those professionals who need post-license in order to give services and to continue their profession.

Continuing education according to Ramachandran (2016), continuing education is any extension of openings for reading, learning and training to professionals that are currently following their completion of or withdrawal from full time schooling. He also emphasized that continuing education is important so that the job aspirants are updated to the current trends of the technology, as the innovation for it is rapidly going off the charts. Also, he said that the continuing education is significant to those professionals that are in the field of fast changing industries.

### Continuing Education in the Philippines

Continuing education in the Philippines can also be described as the other countries describes it, that for every country wants their locally produced professionals to be well-knowledgeable and equipped with the right skills and learnings in relation to their specific filed of expertise, so they help within the community, for their country, and to be globally competitive. But still, there is a difference of acceptance within the group of professionals in the Philippines, in terms of agreeing with the concept of continuing education.

San Luis (2017) mentioned that employers nowadays require, or if not really, prefer professionals that has specific skill set and has qualification to a specific body of knowledge institutions. For example, a project manager who are certified by PMPs.

But in order to achieve this kind of certifications, Filipino professionals must invest for it, meaning, they should pay the price that will deserve the certification. Continuing education is not really an issue for the Filipino professionals, but since mostly of the CPD programs are expensive, it lacks the interest of the professionals to continue education.

### Algorithms used to match professionals with available accredited trainings

### Similar algorithms and their unique differences

Based on the gathered data, the matched technique needed to use is the topic modeling. Topic modeling is a powerful approach for analysis of a large collection of documents. According to Barde et al. (2017), it is used for finding hidden structure from the collection of a document. A topic contains a group of words that often occurs together. Topic modeling can differentiate across the uses of words with dissimilar meanings and link words with the same context. There are 4 methods of topic modeling which includes Latent Semantic Indexing (LSI), Vector Space Model (VSM), Probabilistic Latent Semantic Analysis (PLSA) and Latent Dirichlet Allocation (LDA).

Selivanov (2017) explained the topic modeling as a technique to obtain conceptualize topics from a collection of documents. He discussed the features between the Latent Semantic Analysis and the Latent Dirichlet Allocation. The Latent Semantic Analysis is easy to train, tune and usually work fine in downstream task such as regression, classification, similarity-search and clusterization in a small collection of documents while the Latent Dirichlet Allocation can be quite fast to train in a very large collection of documents.

### Theory and Application of Big Data

Zerhari (2015), big data is typically described by three features called 3Vs (Volume, Velocity and Variety). It concerns to data that are too big, dynamic and difficult. In this framework, data are hard to manage, capture, store and evaluate using old-fashioned data management tools. Therefore, the new conditions executed by Big Data present important tests at different level, comprising data clustering.

Satell (2013) discussed why big data matters to all of us. According to him, big data can generate big brainstorms and very important because it will change how we manage our business. Nowadays, millennials business leaders trusted on "scientific" studies and "statistical significance" to conclude what data could believe. Currently, technology is creating those expectations obsolete and the practice of an organization will never be the similar.

**Use of the selected algorithm in Big Data**

The selected technique in topic modeling for Big Data is the Latent Dirichlet Allocation that can generate statistical model matched to use in this research. Discussed on the research of Maier et. al. (2018), their research goal is to make Latent Dirichlet Allocation topic modelling more available to communication researchers and to guarantee agreement into disciplinary standards. According to them, using LDA to textual data, researchers require to undertake four major challenges that modify these conditions: (a) adequate selection of model parameters, including the number of topics to be generated (b) appropriate pre- processing of the text collection (c) the process of validly interpreting the resulting topics and (d) evaluation of the model's reliability.

Accumulated from the study conducted by Nikolenko et. al. (2014), the study entitled "Measuring Topic Quality in Latent Dirichlet Allocation" explained that Latent Dirichlet Allocation is the modern model of choice for topic modeling. It is an approach to text classification as one document belong to one topic. It also a probabilistic model that can extracts topics from the corpus of documents and the quality of LDA shows better.

**Meta Data Search**

Based on the conducted study of Hemminger B. and Saelim, B. (N.D.), the researchers have usually used bibliographic files to seek out information. Today, the full- text of resources is progressively open for searching, and more researchers are presenting full-text searches. Meaningfully more articles were exposed via full-text searching; however, the accuracy of full-text searching also is pointedly lower than that of metadata searching. Specific features of articles connected with higher significance ratings. By using the number of hits of the search term in the full-text to rank the importance of the article, performance of full-text searching was improved so that both recall, and precision were as good as or better than that for metadata searching. This suggests that full-text searching alone may be sufficient, and that metadata searching as a surrogate is not necessary.

**METHODOLOGY**

**Sources of Data**
The population used in this study comes from the selected different universities around Metro Manila area. The population selected are licensed professionals that are regulated by the PRC. The researcher used purposive sampling to accumulate the data needed for the said study. The researcher selects the professionals that are licensed and those who are continuing their education to answer the given questionnaire. The researcher also used the datasets that contains the list of seminars obtained from the PRC.

**Research Instrument**
The researcher used the questionnaire as the primary instrument in conducting the study which includes the different certain questions. The questionnaire is designed to get the percentage or the total assessment of the PRC services, the evaluation of the benefits of continuing education, and the level of acceptance of the proposed system.

**Respondents' Profile**
The respondents' professional profile and level of agreement in implementing an online CPD application system tool is determined by the evaluation of the following tables, namely: Table 1, 2 and 3.

**Table 1. Length of time as a Professional Identification Cardholder**

|  | Frequency | Percent | Rank |
|---|---|---|---|
| < 1 year | 20 | 20% | 3 |
| 1-3 years | 36 | 36% | 1 |
| 4-6 years | 27 | 27% | 2 |
| 7-10 years | 12 | 12% | 4 |
| more than 10 years | 5 | 5% | 5 |
| Total | 100 | 100% |  |

Based on the accumulated data from the respondents, the highest computed percentage is 36, ranked as number one, which are professional working around 1-3 years, while the lowest computed percentage is 5, ranked as number five, which are those professionals that are working more than years.

**Table 2. In favor of upgrading knowledge and skills as a prerequisite to license renewal**

|  | Frequency | Percent | Rank |
|---|---|---|---|
| Yes | 66 | 66% | 1 |
| No | 34 | 34% | 2 |
| Total | 100 | 100% |  |

Based on the accumulated data from the respondents, the highest computed percentage is 66, ranked as number one, which are professionals that are in favor, while the lowest computed percentage is 34, ranked as number two, which are those professionals that are not in favor.

**Table 3. In favor of renewing their license online if given a chance**

|  | Frequency | Percent | Rank |
|---|---|---|---|
| Yes | 83 | 83% | 1 |
| No | 17 | 17% | 2 |
| Total | 100 | 100% |  |

Based on the accumulated data from the respondents, the highest computed percentage is 83, ranked as number one, which are those in favor of online renewing license, while the lowest computed percentage is 17, which are those who's not in favor of online renewal licensing.

**Data Case Analysis**

The following statistical tools was used to give meaning to the data to be gathered:

1. Frequency and Percentage – this statistical technic was utilized to determine the demographic profile of the respondents of the study.

$$\sum = \frac{fx}{n}$$

Where:
p = percentage
f = frequency
n = total number of respondents

2. Weighted Mean – this statistical technique was used to weigh the answers of the respondents pertaining to every question that was asked on the survey questionnaire except from the respondent's profile.

$$p = \frac{f}{n} \times 100$$

Where:
f = frequency
x = corresponding verbal interpretation
n = total number of respondents

## RESULTS AND DISCUSSIONS

1. Level of Agreement of the Respondents on the Benefits of Continuing Education Thru CPD Law Implementation.

   Based on the evaluation of the respondents, the researcher has analyzed that the professional's diversity between an employer and an employee affected on how they answered the survey. By letting the respondents to freely manipulate the said application, and by analyzing their preferences, the researcher conclude that most of the professionals, both employer and employee agreed on the benefits of continuing education thru CPD law implementation. the highest computed mean is 4.45, ranked as number one, which are employers that agrees on the benefit, while the lowest computed mean is 4.37, ranked as number two, which are those employees that agrees on the benefits. Overall, a total of computed mean 4.40, which defines that both parties agrees on the benefits of continuing education thru CPD law implementation.

2. The respondents' level of agreement on the issues and concerns on the implementation of the CPD Law.

   Based on the accumulated data from the respondents, the highest computed mean is 4.34, ranked as number one, which is the difficulty of affordability, while the lowest computed mean is 4.22, ranked as number four, which is the difficulty of access in far-flung areas. Based on the evaluation of the respondents, the researcher concludes that most of the professionals agreed that the main issue and concern in the implementation of the CPD law is the affordability and accessibility of seminars and other PRC services, like renewal of license.

3. The respondents' level of agreement on the acceptability of the developed solution in terms of Efficiency; Reliability; usability; and Functionality.

   The data gathered in terms of the level of agreement on the acceptability of the developed solution in terms of Efficiency; Reliability; usability; and Functionality, the highest computed mean is 4.50, ranked as number one, which is the functionality, while the lowest computed mean is 4.42, ranked as number four, which is the efficiency. Overall, the computed mean is 4.46, which is agree. Based on the evaluation of the respondents the professionals agree to the use of the proposed application, realizing the benefit they can receive from it.

## CONCLUSIONS AND RECOMMENDATIONS

1. The researcher highly recommends the implementation of the Green Light Application, as it aids the concerns of most of the professionals.
2. The researcher highly recommends having more research about the different applications of the said study.
3. The researcher highly recommends the provision of awareness to the continuing professionals for them to be fully knowledgeable in their current state as a CPD compliant.
4. The researcher highly recommends the use of Latent Dirichlet Allocation algorithm in terms of big data and meta-data mining, as it produce numerous and quality data in just seconds, applicable for mining in documents, articles, and the likes.
5. The researcher recommends the Green Light application developers to improve updates with consideration to security of user access, making it more reliable and robust from malicious intents.
6. The researcher highly recommends the users of the system to study the flow of the system and determine if there are some bugs to fix or improvements to give.
7. The researcher recommends this study and its developed application to be accepted and to be used by the PRC itself.

## REFERENCES

[1] 10 Good Reasons to Continue Your Education (2014), Retrieved on August 2018,
fromhttps://www.chcp.edu/blog/10-good-reasons-continue-your-education
[2] Barde et al. Topic Modeling (2017) Retrieved on September 2018, from
https://ieeexplore.ieee.org/document/8250563
[3] Maier et. al. (2018), Retrieved on September 2018, from
https://www.tandfonline.com/doi/abs/10.1080/19312458.2018.1430754?src=rec sys&journalCode=hcms20
[4] Nikolenko et. al. Measuring Topic Quality in Latent Dirichlet Allocation (2014) Retrieved on September 2018,
from https://logic.pdmi.ras.ru/~sergey/slides/N14_PhMLtalk.pdf
[5] San Luis, G. The importance of continuing education (2017) Retrieved on August 2018, from
http://business.inquirer.net/222641/importance-continuing-education