International Journal of Future Generation Communication and Networking
Vol. 13, No.2s, (2020), pp. 1561–1565

Analysis Of Technology Trends Using Topic Modeling & Question Recommendation System

Rushikesh Jaykar, Roshni Shinde, Kalyani Rane, Pallavi Naikwade and Prof. Varsha khandekar

Department of Information Technology
Savitribai Phule Pune University
Smt, Kashibai Navale college of Engineering
{jaykarrushikesh,shinderoshni04, kalyani2591998, pallavi. Naikwadi27}@ gmail.com

Prof at Department of Information Technology
Savitribai Phule Pune University
Smt, Kashibai Navale college of Engineering
varsha.khandekar@gmail.com

Abstract

Innovation question and answer sites square measure a brilliant stockpile of specialized data. Clients of those sites raise shifted assortments of specialized questions and answer them. These questions cowl a huge change of spaces in designing like Systems, information preparing, Mixed media, Multithreading, web Improvement, Portable Application Advancement, and so on. Breaking down the specific issue substance of those sites will encourage the science and programming framework designing network higher see the needs of engineers and study these patterns in innovation. In that venture, matter information from a prestigious inquiry and answer site known as Stack Flood is dissected abuse the Inert ***Latent Dirichlet Allocation*** *(LDA) theme displaying algorithmic program. The outcomes show that these strategies encourage finds predominant subjects in designer conversations. These subjects square measure broke down to look out some interesting Perceptions like popular innovation/language, the effect of innovation, innovation inclines after some time, a relationship of an innovation/language with elective advancements and examination of advances tending to a region of building or programming framework designing.*

Index Terms – *Multithreading,* ***Latent Dirichlet Allocation,*** *programming framework, Media.*

I. Introduction

The software engineering field has countless innovations. Consistently we are presenting new advancements and they are changing at a quick pace. We proposed an over Profound Intermittent Community oriented Sifting (DRCF) structure with a pairwise positioning capacity for scene suggestion, situated inside six inspired constraints in the cutting edge draws near. Specifically, the proposed structure comprises of two segments (GRMF and MLRP models) that intend to catch both the clients' dynamic and static inclinations from the successions of their registration. Plus, we proposed an oval successive based negative inspecting approach that considers the topographical area of settings to all fly

the cool beginning issue. We study individual and social components identified with weight reduction. To begin with, we played out an underlying examination to comprehend weight reduction with huge scope information from the weight-related portable interpersonal interaction application BOOHER, which is one of the most well known uses of its skin in China. Our point was twofold: right off the bat to find what variables are identified with weight reduction, and afterward to saddle these elements into a weight reduction expectation model. In noting what components are identified with weight reduction we concentrated on separating two kinds of variables, individual and social, and these elements we relater used to develop an important model to anticipate weight reduction consequently.

ISSN: 2233-7857 IJFGCN
Copyright ⓒ2020 SERSC

we propose a start to finish discourse language identification(SLD) model. Three measures are utilized to cause the prepared model to can appropriate for short expressions. In the waveform area, we utilize a period scale modification(TSM) technique to expand the length of info articulations. In the component space, we utilize the exchange learning thought to prepare a profound phoneme classifier, bottleneck highlights of the phoneme classifier which incorporate phoneme discriminative data are utilized to prepare language classifiers. In the language classifier area, a LSTM base classifier is prepared by brief timeframe highlight squares which can make the prepared model fitting for brief term inputs. we introduced the Community oriented Denoising AutoEncoder (CDAE) for the top-N suggestion issue. CDAE learns appropriated portrayals of the clients and things through detailing the client thing criticism information utilizing a DE noising Auto-Encoder structure. A few past works can be viewed as extraordinary instances of the proposed model. We led a far reaching set of analyses on three informational indexes to concentrate how the decision of the model parts impacts the presentation. The changed component factors are largely beta dispersed. With the separation connection metric, the decorrelation execution of the proposed nonlinear change was shown to be better than those of PCA and ICA with both orchestrated and genuine information. Additionally, we applied the proposed nonlinear change in two applications, i.e., quantization of line ghastly recurrence parameters in the discourse straight prescient model and EEG signal classification. Broad test results demonstrated that, when doing decorrelation and highlight choice for unbiased like information, the proposed equal nonlinear change (PNT)- based nonlinear change can accomplish better useful execution and is desirable over the expectedly applied PCA-based straight change.

II. TOPIC MODELLING

1. LDA

AI pack systems don't appear to be the sole gratitude to remove subjects from a book information set. Content mining writing has arranged many applied math models, alluded to as probabilistic theme models, to watch subjects from partner degree untagged arrangement of archives. one among the premier popular models is the dormant Dirichlet allotment (LDA) calculation created by Blei, Ng, and Jordan [i]. LDA could be a generative unaided probabilistic algorithmic program that disconnects the top K subjects during an information set as portrayed by the first applicable N catchphrases. In various words, the archives inside the information set region unit drawn as arbitrary blends of idle points, any place each theme is described by a Dirichlet distribution over an immovable jargon. "Inactive" signifies subjects should be derived rather than legitimately found. The algorithmic program is sketched out as a generative model [ii], which infers that it relies upon approximately from the earlier applied math assumptions [iii], i.e.: Word request in archives isn't vital. Archive request inside the information set isn't vital. The quantity of subjects must be most popular prior. A similar word will have a place with different subjects. Each record inside the absolute variety of D archives is viewed as a mix of K inert points. Every point contains a multinomial appropriation over a jargon of words w.
The generative process for LDA is given by:
$\theta j \sim D[\alpha]$, $\Phi k \sim D[\beta]$, $zij \sim \theta j$, $xij \sim \Phi zij$
where:
θj is the mixing proportion of topics for document j, and it is modeled by a Dirichlet distribution with parameter α.
A Dirichlet prior with parameter β is placed on the word-topic distributions Φk.
zij= k is the k topic drawn for the ith word in document j with probability θk| j.
Word xij is drawn from topic zij, with xij taking on value w with probability Φw|zij.

ISSN: 2233-7857 IJFGCN
Copyright ©2020 SERSC

Let's have a look at the learning phase (or posterior computation). We have selected a fixed number K of topics to discover, and we want to learn the topic representation for each document and the words associated with each topic. Without going into too much detail, the algorithm operates as follows:
At first, the algorithm randomly assigns each word in each document to one of the K topics. This assignment produces an initial topic representation for all documents and an initial word distribution for all K topics.
Then, for each word w in document d and each topic t, the algorithm computes:

p(topic t | document d), which is equal to the proportion of words in document d currently assigned to topic t.
p(word w | topic t) which corresponds to the proportion of assignments to topic t over all documents containing this word w.
Reassign word w to a new topic t, based on probability p(topic t | document d) * p(word w | topic t).
It then repeats the previous two steps iteratively until convergence.
There are several variations and extensions of the LDA algorithm, some of them focusing on better algorithm performance.

2.NEURAL NETWORK TOPIC MODELS

Theme models are a class of probabilistic models for content examination, broadly utilized in many research territories that utilization printed information as their exploration material, for example writing contemplates, history of thoughts, social and political theory, and necessities building. The most broadly utilized point model, Idle Dirichlet Examination (LDA) [Blei et al., 2003] is a various leveled Bayesian model that is regularly actualized utilizing MCMC or variety induction techniques.

As of late, a few techniques have been conceived to define progressive Bayesian models utilizing target capacities to be limited utilizing angle based strategies [Kingma and Welling, 2014]. This basically brings Bayesian models, for example, LDA into the neural system world [Srivastava and Sutton, 2017; Miao et al., 2017]. This permits the parameters of the models to be evaluated utilizing proficient equipment and programming ordinarily used to prepare neural models, however maybe more strikingly it permits us to detail more fascinating models that are less oversimplified than LDA or that can work with new sorts of information, not solely message.

We can see a few bearings for an Ace's proposition venture around neural system based point models and we are available to conversation about the exploration question the undertaking would examine. For example, the task could be a cautious examination of the upsides and downsides of utilizing neural surmising contrasted with customary induction techniques, exploring perspectives, for example, preparing speed, soundness, theme quality, and interpretability. Or then again you could characterize another subject model for another sort of information, for example, pictures or diagram organized information, by joining the neural point model with a neural system for your favored kind of information.

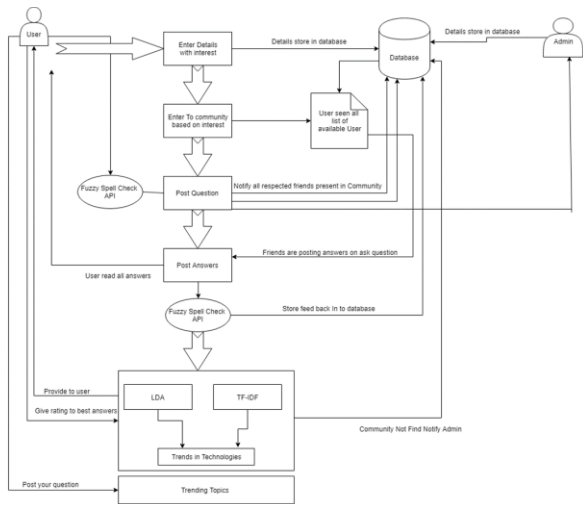
Your first task will be to reimplement a previously published model, and this will then be your starting point for further extensions and evaluations.

III. OVERALL ARCHITECTURE DIAGRAM

We are going to find the current trending topic based on Text. There are five main steps involved in the implementation: data extraction, data pre-processing, topic modeling, post-processing, inferring results and

Copyright ©2020 SERSC

creating a visualization of trends. Topic modeling using the LDA algorithm to analyze the trends in technologies and languages. From text corpus we need to extract topics, it will be done performing NLP on the text corpus. the question is searched by the user. The answer list will be provided to the user. As the machine is already trained previously by the stack overflow dataset. Based on the question and answer the trending topic will be detected. The LDA algorithm going to apply on the QA corpus to find a trending topic. After getting the current trending topic study material will be provided to the user.



CONCLUSION

Right now, theme following exploration is directed dependent on point model. Right off the bat, the LDA model is utilized to extricate the point data from the news writings of various time windows. At that point the improved Single-Pass calculation is utilized for subject following, in which the time rot work and the JS dissimilarity are utilized to quantify the closeness between the themes. At last, for the consequences of point following, the substance

REFERENCES

[1] J. Manotumruksa, C. MacDonald, and I. Ounis, ``A deep recurrent collaborative_ltering framework for venue recommendation," in *Proc. 26th ACM Int. Conf. Inf. Knowl. Manage.*, Singapore, Nov. 2017, pp. 1429_1438.

[2] Z. Wang, T. Derr, D. Yin, and J. Tang, ``Understanding and predicting weight loss with mobile social networking data," in *Proc. 26th ACM Int. Conf. Inf. Knowl. Manage.*, Singapore, Nov. 2017, pp. 1269_1278.

[3] Z. Ma, H. Yu, W. Chen, and J. Guo, ``Short utterance based speech-language identi_cation in intelligent vehicles with time-scale modi_cations and deep bottleneck features," *IEEE Trans. Veh. Technol.*, to be published, DOI: 10.1109/TVT.2018.2879361.

[4] Y. Wu, C. DuBois, A. X. Zheng, and M. Ester, ``Collaborative denoising auto-encoders for top-n recommender systems," in *Proc. 9th ACM Int. Conf. Web Search Data Mining*, San Francisco, CA, USA, Feb. 2016, pp. 153_162.

Copyright ⓒ2020 SERSC

[5] Z. Ma, J.-H. Xue, A. Leijon, Z.-H. Tan, Z. Yang, and J. Guo, ``Decorrelation of neutral vector variables: Theory and applications," *IEEE Trans.Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 129_143, Jan. 2018.

[6].Wang, W. Lu, M. Ester, C.Wang, and C. Chen, ``Social recommendation with strong and weak ties," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, Indianapolis, IN, USA, Oct. 2016, pp. 5_14.

[7] H. Huang, Q. Zhang, J. Wu, and X. Huang, ``Predicting which topics you will join in the future on social media," in *Proc. 40th Int. ACMSIGIR Conf. Res. Develop. Inf. Retr.*, Tokyo, Japan, Aug. 2017, pp. 733_742.

[8] X. Wang, X. He, L. Nie, and T.-S. Chua, ``Item silk road: Recommending items from information domains to social users," in *Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Tokyo, Japan, Aug. 2017, pp. 185_194.

[9] X. Yang *et al.*, ``Collaborative _ltering-based recommendation of online social voting," *IEEE Trans. Comput. Social Syst.*, vol. 4, no. 1, pp. 1_13, Mar. 2017.

[10] P. Li, Z. Wang, Z. Ren, L. Bing, and W. Lam, ``Neural rating regression with abstractive tips generation for the recommendation," in *Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Tokyo, Japan, Aug. 2017, pp. 345_354.

Copyright ⓒ2020 SERSC