Scalable Algorithms For Cqa Post Voting Prediction

MICANS INFOTECH

ABSTRACT

- Community Question Answering (CQA) sites, such as Stack Overflow and Yahoo! Answers, have become very popular in recent years.

- These sites contain rich crowdsourcing knowledge contributed by the site users in the form of questions and answers, and these questions and answers can satisfy the information needs of more users.

- In this article, we aim at predicting the voting scores of questions/answers shortly after they are posted in the CQA sites.

- To accomplish this task, we identify three key aspects that matter with the voting of a post, i.e., the non-linear relationships between features and output, the question and answer coupling, and the dynamic fashion of data arrivals.

EXISTING SYSTEM

- In this article, we aim at predicting the voting scores of questions/answers shortly after they are posted in the CQA sites.

- To accomplish this task, we identify three key aspects that matter with the voting of a post, i.e., the non-linear relationships between features and output, the question and answer coupling, and the dynamic fashion of data arrivals. A family of algorithms are proposed to model the above three key aspects.

- Generally speaking, there are three key aspects that matter with the voting prediction of a post, namely, (1) the non-linearity between features and output, (2) the coupling between questions and answers, and (3) the dynamics (of training data sets).

- First, both the contextual features (e.g.,the reputation of the user who issues the question, etc.) andthe content of the post

DISADAVANTAGE

- The non-linear relationships between features and output, the question and answer coupling, and the dynamic fashion of data arrivals.

- Lists the main symbols we use throughout the paper. For convenience, we use bold capital letters for existing matrices/vectors at time t, and bold lower case letters for newly arrived matrices/vectors at time t + 1.

- Let us first define the solution space of the lip problem, which is represented by a genealogy graph in f in this section, we propose our solutions for the lip problem.

- We start with presenting two algorithms for problem 1 (subsection 3.1) and problem 2 (subsection 3.2), respectively; and then address the computational challenges (subsection 3.3-3.4)

PROPOSE SYSTEM

- Some of the proposed algorithms (lip-kim, lip-kima, and lip-kimaa) can capture three key aspects (non-linearity, coupling, and dynamics) that matter with the voting score of a post

- In this section, we propose our solutions for the LIP problem. We start with presenting two algorithms for Problem 1 (subsection 3.1) and Problem 2 (subsection 3.2), respectively; and then address the computational challenges

- (subsection 3.3-3.4).

- For the coupling aspect, LIP-KM first transfers the features and then imposes a so-called voting consistency on the prediction space by requiring the predicted voting of a question to be close to that of its answer (see our techreport [3] for the detailed explanations about its rationality).

ADVANTAGES

- 3.4 LIP-KIMAA algorithm compared with LIP-KIM, LIP-KIMA is much more scalable, being linear in terms of both time and space complexity

- If only a subset of these three aspects matter with the prediction performance for some applications,

- Our algorithms can be naturally adapted to these special cases (e.G., See LIP-IM and LIP-KIA in the supplementary material).

- Here, we further show the flexibility of our models in terms of temporal forgetting and feature/example selection.

SOFTWARE REQUIREMENT

- **HARDWARE REQUIREMENT:**
- System   :   Pentium IV 2.4 GHz.
- Hard Disk   :   40 GB.
- Floppy Drive   :   1.44 Mb.
- Monitor   :   15 VGA Colour.
- Mouse   :   Sony.
- Ram   :   512 Mb.

Cont..

- **SOFTWARE REQUIREMENT:**
- Operating system : Windows XP.
- Coding Language : ASP. Net with C#
- Data Base : SQL Server 2005.

CONCLUSION

- In this aricle, we have proposed a family of algorithms to comprehensively and efficiently predict the voting scores of questions/answers in cqa sites. In particular,

- Some of the proposed algorithms (lip-kim, lip-kima, and lip-kimaa) can capture three key aspects (non-linearity, coupling, and dynamics) that matter with the voting score of a post,

- While others can handle the special cases when only a fraction of the three aspects are prominent.