Scalable Algorithms for CQA Post Voting Prediction

Abstract:

Community Question Answering (CQA) sites, such as Stack Overflow1 and Yahoo! Answers2, have become very popular in recent years. These sites contain rich crowd sourcing knowledge contributed by the site users in the form of questions and answers, and these questions and answers can potentially satisfy the information needs of more users. For example, millions of programmers ask and answer questions on Stack Overflow, and even more users now use Stack Overflow to seek solutions for their programming problems.

Existing System:

We focus on the voting score prediction of questions/answers shortly after they are posted in the CQA sites. Such a task is essential for the prosperity and sustainability of the CQA ecosystem, and it may benefit all types of users, including the information producers and consumers. For example, detecting potentially high score answers can benefit the questioners as well as the people who have similar questions; it would also be helpful to identify high-score questions in the early stage and route them to expert answerers.

Proposed System:

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. In terms of computation efficiency, some algorithms (LIP-IM, LIP-IMF, LIP-KIA, LIP-KIMAA, and LIP-KIMAA) enjoy linear, sub-linear, or even constant scalability. The proposed algorithms are also able to fade the effects of old examples (LIP-IMF), and select a subset of features/examples (LIPMS and LIP-KMS). We analyze our algorithms in terms of optimality, correctness, and complexity, and

reveal the intrinsic relationships among different algorithms. We conduct extensive experimental evaluations on two real data sets to demonstrate the effectiveness and efficiency of our approaches.

Advantages:

- More Effectiveness.
- More Efficiency.

Modules:

- LIP-KM Algorithm.
- LIP-KIMA Algorithm
- LIP-KIMAA Algorithm

SYSTEM REQUIREMENTS

H/W System Configuration:-

Processor	-	Pentium –III
RAM	-	256 MB (min)
Hard Disk	-	20 GB
Key Board	-	Standard Windows Keyboard
Mouse	-	Two or Three Button Mouse
Monitor	-	SVGA

S/W System Configuration:-

Operating System	:	Windows95/98/2000/XP
Application Server	:	Tomcat5.0/6.X
Front End	:	HTML, Jsp

Scripts	:	JavaScript.
Server side Script	:	Java Server Pages.
Database	:	MySQL 5.0
Database Connectivity	:	JDBC