

Title: Data-Driven Intelligence: Challenges and our Solutions

Abstract: Data-Driven Intelligence has become a hot research topic in the area of information science. This talk aims to outline the challenges on Data-Driven Intelligence. Then our solutions for Data-Driven Intelligence are provided, which cover the following aspects. 1) A hierarchical entropy-based approach is demonstrated to evaluate the effectiveness of data collection, the first step of Data-Driven Intelligence. 2) A multi-view-based method is illustrated for filling missing data, the preprocessing step for Data-Driven Intelligence. 3) A unified framework is outlined for Parallel Large-scale Feature Selection to manage Big Data with high dimension. 4) A MapReduce-based parallel method together with three parallel strategies are presented to compute rough set approximations for classification, which is a fundamental part in rough set-based data analysis similar to frequent pattern mining in association rules. 5) Incremental learning-based approaches are shown for updating approximations and knowledge in dynamic data environments, e.g., the variation of objects, attributes or attribute values, which improve the computational efficiency by using previously acquired learning results to facilitate knowledge maintenance without re-implementing the original data mining algorithm. 6) A deep-learning-based model to deal with multiple different sources of data is developed. 7) Several applications on high speed train and urban computing, etc. are shown.