

Feature Selection Combining Redundant Information and Prior Knowledge

- Proposed a Mixed-integer model which can select novel features which contain less redundancy and fit well with prior knowledge.
- Digged into the structure of the optimal solutions. Provided several ways to speed up the solving process. These speed-up methods worked well especially with big data.
- Theoretically analyzed the NP-hard problem and proved that this problem is polynomial-solvable under certain conditions.
- The existing experimental results showed better performances over traditional feature selection methods in the machine learning area. **[In progress]**

Contemporaneous Health Index

- Developed a novel formulation for contemporaneous patient risk monitoring. This formula translated multivariate longitudinal measurements into a contemporaneous health index (CHI) that captures patient condition changes over the course of progression. The formulation can work in both supervised and unsupervised manner.
- Proposed algorithms to mitigate the challenges associated with the nonsmooth convex optimization problem. Our algorithms involved dual reformulation and block coordinate descent. Extensive numerical studies were performed on both synthetic datasets and real-world applications on Alzheimer's disease and Surgical Site Infection.

Interesting Course Projects Related to Machine Learning

- **Sparse Principle Component Analysis (sPCA)**. Proposed an algorithm combining alternating maximization (AM) method and convex optimization to solve the NP-hard problem. Only a few explained variances were sacrificed and the orthogonality of PCs were guaranteed.
- **Training neural networks with Particle Swarm optimization (PSO)**. A global optimization method called PSO was used to search for the optimal parameters of a neural network. Its performances were compared with the traditional gradient methods.
- **K kernel nearest-neighbor algorithm for the Netflix competition**. Implemented kNN, Kernal kNN, weighted kNN for the Netflix competition. k was found by leave-one-out cross validation.
- **Scaling up K-Means with MapReduce**. Implemented the k-means algorithm in Hadoop MapReduce to cluster the BBC News with the TF-IDF values.
- **PageRank with MapReduce**. Implemented the multiplication of big matrix with MapReduce to find the stationary distribution of the transition matrix.

Interesting Course Projects Related to Reinforcement Learning

- **Cart-pole problem**. Implemented policy-based reinforcement learning to solve for optimal movements which balance a pole connected with one joint on top of a moving cart. The performances of two learning methods (Monte-Carlo policy gradient and Actor-Critic policy gradient) are compared.
- **Kalah game**. Used minimax search with alpha-beta pruning as the basic algorithm. Proposed various heuristic functions and created different AIs to beat the basic AI.

A Vision-based Intelligent Off-road Robot

- The leader and programmer of the team. Developed a real-time object recognition system (based on OpenCV) which navigated the robot by detecting markers from the messy backgrounds. The system worked perfectly in the changeable out-door environment under various backgrounds and weather conditions.
- Designed the mechanical control system which could accurately complete the action of fetching a certain item and the action of climbing stairs.