CS670/470 Team Project Phase 2: Online Feature Selection

UMass-Boston 11/16/2017

1 Educational Goal

Practice how to leverage online learning algorithms for feature selection.

2 Details

Project goal: Implement the Scalable and Accurate OnLine Approach (SAOLA) to learn how to apply online learning to feature selection.
Due Date: 4:00 pm, November 28, 2017
Programming language: Python 2.7.

3 Tasks

3.1 Task 1

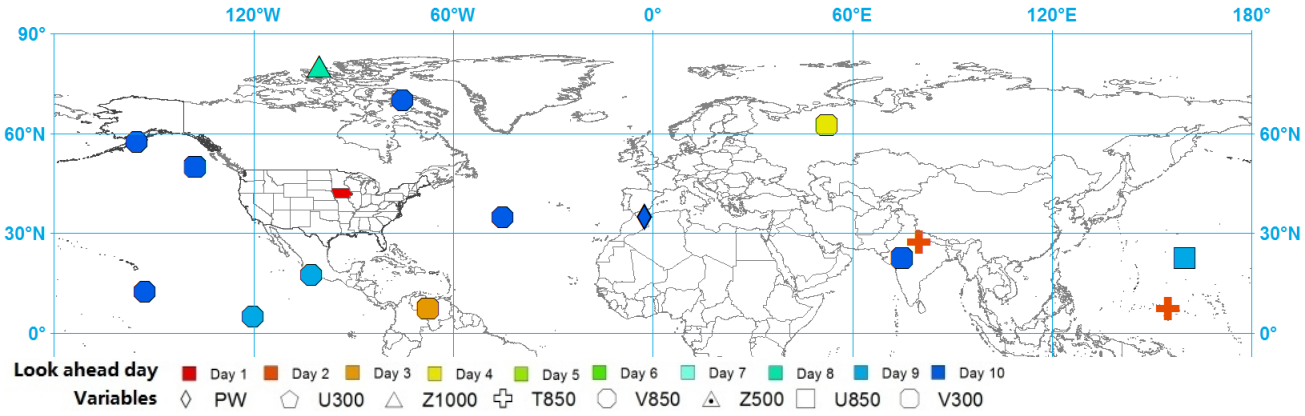
Review the lecture notes in "CS670_Team_Project_Part1.pdf" and implement the Scalable and Accurate OnLine Approach to feature selection (SAOLA). Apply it to the experimental samples which you created in Phase 1 to get the most relevant features (Markov Blanket). To speed up the calculations, you can simplify the task by applying the SAOLA algorithm to the last 4 days of features $\{m_{t-3}^S, m_{t-2}^S, m_{t-1}^S, m_t^S\}$ instead of the entire 10 days of features that you compiled in phase 1.

3.2 Task 2 (Optional) Bonus points

Modify SAOLA to improve its performance on unbalanced data.

3.3 Task 3 (Optional) Bonus points

Highlight the relevant features you got on a map.
A example solution of task 3



4 Submission Requirements

Submit the code and a report (include the important features you identified in task 1. If you finish task 2 and task 3, please submit a summary of your improvement for task 2 and your map image for task 3) through your UMassOnline account. Each team can submit a single solution through the team lead's UMassOnline account; the submission should include a list of all team members' names.