Chapter 9

Feature Selection Based on Elements of Game Theory

FEATURE SELECTION BASED ON THE SHAPLEY VALUE

This chapter describes an approach to feature selection originating from the principles of game theory, in particular, in the context of coalition games (Cohen, Ruppin, & Dror, 2005), (Cohen, Dror, & Ruppin, Feature selection via coalition game theory, 2007). Feature selection is iteratively done through the Contribution-Selection algorithm aiming at optimizing classifier performance on unseen data. This algorithm combines both filter and wrapper approaches to feature selection. In contrast to the conventional filters, features are iteratively re-ranked by a classifier. The ranking itself employs the Shapley value in order to estimate the importance of each feature for the given task.

The algorithm assumes that two sets are available: training set and validation set. The former is employed for training a classifier while the latter is used to

DOI: 10.4018/978-1-60960-557-5.ch009
evaluate the trained classifier performance, given a subset of features - candidates for selection. The classifier performance is measured by classification accuracy, though, in principle, any other suitable characteristic, such as the area under the Receiver Operating Characteristic (ROC) curve or False Acceptance Rate (FAR), can be utilized as well.

**Main Idea**

Coalition games belonging to cooperative game theory involve a set of players and a reward associated with different groups or coalitions of players. The reward of a certain coalition depends on individual contributions of players composing this coalition to the game. The larger the contribution of a player is, the higher the benefit of having this player in a coalition. Coalitions with high reward are naturally preferable over those with small reward.

The contribution is based on the Shapley value (Shapley, 1953). This value is defined as follows. Let the marginal importance of player \( i \) to a coalition \( S \ (i \not\in S) \) be

\[
\Delta_i(S) = v(S \cup \{i\}) - v(S),
\]

where \( v(S) \) is the reward associated with coalition \( S \). The reward can be negative, zero, or positive. The negative or zero reward implies no benefits of inclusion of player \( i \) into the current coalition.

The Shapley value is then defined as

\[
\Phi_i(v) = \frac{1}{n!} \sum_{\pi \in \Pi} \Delta_i(S_i(\pi)),
\]

where \( n \) is the total number of players, \( \Pi \) is the set of permutations over \( n \), and \( S_i(\pi) \) is the set of players appearing before player \( i \) in permutation \( \pi \). Thus, the Shapley value of a given player is the mean of its marginal importance averaged over all possible coalitions of players.

Casting these game theory definitions into feature selection terminology maps the set of players into the set of features and the reward given to a coalition of players into the value of the classifier performance characteristic (here, the accuracy), attained with a certain subset of features. From now on, we will utilize feature selection terminology instead of game theory one.
Related Content

A Comparative Study of an Unsupervised Word Sense Disambiguation Approach
[www.igi-global.com/chapter/comparative-study-unsupervised-word-sense/76119?camid=4v1a](www.igi-global.com/chapter/comparative-study-unsupervised-word-sense/76119?camid=4v1a)

MicroRNA Precursor Prediction Using SVM with RNA Pairing Continuity Feature
Huan Yang, Yan Wang, Trupti Joshi, Dong Xu, Shoupeng Yu and Yanchun Liang (2011). *Interdisciplinary Research and Applications in Bioinformatics, Computational Biology, and Environmental Sciences* (pp. 73-82).
[www.igi-global.com/chapter/microrna-precursor-prediction-using-svm/48366?camid=4v1a](www.igi-global.com/chapter/microrna-precursor-prediction-using-svm/48366?camid=4v1a)
State-of-the-Art Neural Networks Applications in Biology
[www.igi-global.com/article/state-of-the-art-neural-networks-applications-in-biology/105598?camid=4v1a](www.igi-global.com/article/state-of-the-art-neural-networks-applications-in-biology/105598?camid=4v1a)

The Fuzzy Integrated Energy Prior-warning Model Based on Entropy Weight
Yaqun He, Hua Wei, Weiran Zuo, Xiaobing Wu, Xin Ge, Shan Wu and Baofeng Wen (2011). *Interdisciplinary Research and Applications in Bioinformatics, Computational Biology, and Environmental Sciences* (pp. 289-295).
[www.igi-global.com/chapter/fuzzy-integrated-energy-prior-warning/48385?camid=4v1a](www.igi-global.com/chapter/fuzzy-integrated-energy-prior-warning/48385?camid=4v1a)