Multi Class Support Vector Machine

Dr. Kalidas Y., IIT Tirupati

Opinion vector

- Consider given an image, we have to say, if it is cat or not-cat
  - Is it classification or regression problem?
  - How many classes?
- Let us say there are 5 experts
- Given x, each person says how much it looks like a cat
  - What is x corresponding to?
- Can we treat each expert as an ‘abstract sensor’?
- Can we construct a 5 dimensional vector?

- If there are 100 experts, then what is the size of the vector constructed?
- What are your thoughts about this ‘abstract sensor’?
  - Can they be humans?
  - Can they be other classifiers?
  - Can there be sensor-having-sensors-having-sensors-having-sensors.. kind of very nested structure?
  - Can a **feature vector** be engineered in this fashion? Is this one way of **feature engineering**?

97) key phrase… “Multi class SVM”

98) key phrase… “One vs Rest SVM”

- Output yi is a k-class data point
- STEP 1: Build “k” SVMs
- STEP 2: For each class i, build SVM such that Positives=class i, Negatives=Others
- STEP 3: For each data point compute class predictions
  - xnew is input
  - SVM_1(xnew), SVM_2(xnew),…,SVM_k(xnew)
  - Treat this as data transformation
  - xnew’ = [.. .. ..] (above predictions, where each value is +1 or -1)
- STEP 4 :
  - Build Multi Class Logistic Regression
  - Transform each data point in Data Set D, and created D’ (each point is k dimensional)
  - On the D’ use Multi Class Logistic Regression

99) key phrase… “One vs One SVM”

- Output yi is a k-class data point
- STEP 1: Build “k*(k-1)/2” SVMs
- STEP 2: For each pair of classes i and j, build an SVM_ij such that Positives=class i, Negatives=class j
- STEP 3: For each data point compute class predictions
  - xnew is input
  - SVM_11(xnew), SVM_12(xnew),…,SVM_k-1,k(xnew) //k*(k-1)/2 such SVMs are there
  - Treat this as data transformation
  - xnew’ = [.. .. ..] (above predictions, where each value is +1 or -1)
- STEP 4 :
  - Build Multi Class Logistic Regression
  - Transform each data point in Data Set D, and created D’ (each point is k*(k-1)/2 dimensional)
  - On the D’ use Multi Class Logistic Regression