Committees of Classifiers

- □ Classifier Committee (CCs) is based on applying k different classifiers h_1, .. , h_k to the same task and then combining their outcomes
- □ Usually the classifiers are chosen to be different in some respect
  - Different indexing approach
  - Different learning method applied
  - Different types of errors !!
- □ It must be defined a way to combine them
- □ Justified only by superior effectiveness

Combination rules

- □ Majority Voting: the classification decision that reach the majority of votes is taken
- □ Weighted Linear Combination: a weighted sum of the k CSV_i's yields the final CSV_i
- □ Dynamic Classifier Selection: the judgment of the classifier h_t that yields the best effectiveness on the validation examples most similar to d_j is adopted
- □ Adaptive Classifier Combination: the judgment of all the classifiers are summed together, but their individual contribution is weighted by their effectiveness on the examples most similar to d_j

Boosting

- Boosting is a CC method whereby the classifiers ('weak hypothesis') are trained sequentially by the same learner ('weak learner'), and are combined into a CC ('final hypothesys')
- The training of h_t is done in such a way to try to make the classifier to perform well on examples in which $h_1,..,h_{t-1}$ have performed worst
- AdaBoost is a popular Boosting algorithm

Freund & Schapire's AdaBoost

At iteration s:

1. Passes a distribution D_s of weights to the weak learner, where $D_s(d_j)$ measures how effective $h_1,..,h_{s-1}$ have been in classifying d_j
2. The weak learner returns a new weak hypothesis h_s that concentrates on documents with the highest D_s values
3. Runs h_s on Tr and uses the results to produce an updated distribution D_{s+1} where
   - Correctly classified documents have their weights decreased
   - Misclassified documents have their weights increased

Evaluating TC systems

- Similarly to IR systems, the evaluation of TC systems is to be conducted experimentally, rather than analytically
- Several criteria of quality:
  - Training-Time efficiency
  - Classification-Time efficiency
  - Effectiveness
- In operational situations, all three criteria must be considered, and the right tradeoff between them depends on the application

Types of predictions [Aiolli05]

- Ordering Predictions
  - Ordering of classes (or documents) on a relevance basis in such a way to be consistent with the supervision given as partial orders over the classes (or documents)
  - Single-label classification, ranking
- Rating Predictions
  - Giving ranks from an ordinal scale to examples
  - Binary classification, ordinal regression, and their multivariate extensions

Supervision

- Supervision can be described as conjunctive sets of preferences of two types
  - Qualitative Preferences
    - $(u(d_i,y_r), u(d_j,y_s))$
  - Quantitative Preferences $(\tau \in \mathcal{R})$
    - $(u(d,y), \tau)$
    - $(\tau, u(d,y))$

Linear Preferences

- Now, consider linear expansion of the relevance function
  - $u(d,y) = w \cdot \phi(d,y)$
  - where $\phi(d,y) \in \mathcal{R}^d$ is a joint representation of document-class pairs and w a weight vector
- Qualitative preferences can be written as
  $w \cdot (\phi(d_i,y_r)- \phi(d_j,y_s)) > 0$
- Quantitative preferences $\delta(u(d,y)-\tau)$ can be written as
  $(w,\tau) \cdot (\delta\ \phi(d,y),-\delta)>0$

Summarizing

- □ All the problem setting above can be seen as homogeneous linear problems in an opportune augmented space
- □ Any algorithm for linear optimization (e.g. perceptron, SVMs, or a linear programming package) can be used to solve them

Examples of topics.. Applications of IR

- Multimedia IR, Video and image retrieval, Audio and speech retrieval, Music retrieval
- Question answering, Summarization
- Cross-language retrieval, Multilingual retrieval, Machine, translation for IR
- Interactive IR (User interfaces and visualization, User studies, User models, Task-based IR, User/Task-based IR theory)
- Web IR, Intranet/enterprise search, Citation and link analysis, Adversarial IR

Examples of topics.. Techniques

- Web Crawling
- Information Extraction, Lexical acquisition
- XML and metadata retrieval, Ontology learning
- NLP processing, Language Models for IR
- Automatic building of Thesauri

Examples of topics..
Advanced ML techniques

- Feature Selection
- Similarity metrics and (structured) Kernels for IR
- Preference Learning and atypical tasks in IR
- Semi-supervised Learning for text categorization and ranking
- Classification on Structured output (e.g. hierarchies)

Main Conferences and Journals

- Conferences
  - ACM SIGIR, Special interest group on IR
  - CIKM, Conference on Information and Knowledge Management
  - ECIR, European Conference on IR
- Journals
  - Information Retrieval (Springer)
  - Information Research (Electronic)