MCS 2004 Round Table Discussion

Thursday, June 10th

Moderators: Joydeep Ghosh and Nikunj Oza

Story telling “competition (with prize!)”

- Three broad approaches to MCS:
  - combining strong learners
  - combining weak learners
  - modular nets and hybrids

- Each approach has some theoretical backing, many specific methods/solutions
  - Strong (bagging, ensembles,…)
    - Variance reduction
    - Ensemble error = average error – average “mismatch”
    - Coverage and decision optimizations
  - Weak (arcing/boosting, random forests,…)
    - Reduces bias and variance
    - Gradient descent in function space; Additive logistic regression
    - Adaboost; logitboost, stochastic gradient boosting,..
  - Modular:
    - divide-and-conquer, complementary data sources

Stories

- Share an experience/anecdote showing a remarkable success/failure in using a specific approach
  - contrasts with the result using a different approach
  - provides some insight into best practice or malpractice.

  - Success stories of hybrid systems excluded
  - Everyone rates stories; winner announced on Friday.

  - story could be completely fabricated,
    - Audience: spot this and argue why the story is phony

Aim: approaches ←→ data characteristics.

Stories

- Larry Hall: DTs > SVMs, ensembles of SVMs, NNs on Forest Covertype.
- Piero 1: Insurance underwriting, Random Forest 1% better than SVMs, MARS, others. Majority vote 0.2% better than RF. RFs simple and great!
- Piero 2: Predictions. Attribute selection, weighting. Used Evolutionary Algorithm. Efficient. Applied to vehicle fleet selection.
- Rainer:
  - 70-dimension feature vector divided into 4 to 8 dims, 14 classifiers, sum rule, worse than best individual (range 80-90%). Combiner in between best and worst.
- Horst (**16**): MADE UP (mostly)! THIEF CONVICTED THANKS TO MCSs. Sports car wheels stolen. Owner sees wheels on old car. Matched polish on wheels to car polish. MCSs used for diatome classification to perform match. Thief says, “I hate MCSs!!”

Stories 2

- Josef: Face recognition. Combine classifiers on each color channel. Performance up 40%. Used single classifier in concatenated feature space.
- Nagi Rao (**18**): Manufacturing problem. Classify, localize objects in manufacturing cell. Used laser rangefinder on transparent objects. But used those features as input to MCSs. Perceptron worked.
- Phil Kegelmeyer (**15**): Strong (human) learners vs. weak. Evaluating optics. Three humans. Random forests on bad features > three humans (are they useless?).

Looking to the Future

- Topics to de-emphasize
  - Already resolved by some definitive works
  - little chance of further progress

  - **Strong:** (1) [**7**] new diversity measures; (2) [**4**] new data independent combinations,..
  - **Weak:** (3) [**8**] new boosting variants,….
  - **Multiclass:** (4) [**8**] new ECOC codes;,..
  - **Theory:** (5) [**0**] math quantification of gains
    - Meta-theory: given data properties (size, dimensionality, complexity measures), prescribe best MCS approach and predict ensemble gains
  - (6). [**1**] OVA vs ECOC vs. Round-robin
  - (7). [**3**]Really no more boosting variants!!!

Directions for MCS

- Apply to fully/partially unlabelled data
  - Cluster ensembles
- on-line applications; streaming data
  - highly non-stationary environments

- applying to new domains with complex data forms
  - relational tables
  - web data (text and links)
- distributed data mining

  - Implications for feature selection

Directions for MCS

- 1. [11] Apply to fully/partially unlabelled data (Ayad)
  - Cluster ensembles
- 2. [11] on-line applications; streaming data (Kuncheva)
- 3. [7] highly non-stationary environments (", Polikar)
  - Many forms: noise, seasonal variations, etc.
- 4. [0] Ensembles for recommenders, reinforcement learning.
- 5. [9] applying to new domains with complex data forms
  - e.g., biological data
- 6. [7] distributed data mining
- ************************
- 6.5. [1] Life cycle (Bonissone).
- 7. [3] Missing values (Melville, Duin)
- 8. [4] Structural components (Bunke)
- 9. [2] Sensor Fusion (Rao)
- 10. [5] Multi-stage approaches to multi-class, with tailored feature selection (Sansone, Ghosh)
- 11. [5] Accuracy vs. ease of maintenance.