Train In Data

Bayes Conditional Risk

Cost Sensitive Approaches

- CSL
  - Direct approaches
    - Misclassification cost into the training of the classifier
  - Meta-learning
    - Pre-processing
      - Under-sampling
      - Over-sampling
    - Post-processing
      - Modify the outputs of the classifier

Cost Matrix

C(i,j) ➔ cost of assigning an observation of class j to class i

C(0,0) and C(1,1) ➔ Cost of correct classification, usually 0

C(0,1) and C(1,0) ➔ Cost of FN and FP, respectively, usually 1

	Real Negative	**Real Positive**
Predicted Negative	C(0,0) TN	C(0,1) FN
Predicted Positive	C(1,0) FP	C(1,1) TP

Cost Matrix

C(i,j) ➔ cost of assigning an observation of class j to class i

C(0,0) and C(1,1) ➔ Cost of correct classification, usually 0

C(0,1) and C(1,0) ➔ Cost of FN and FP, respectively, usually 1

	Real Negative		Real Positive	
Predicted Negative	C(0,0)	0	C(0,1)	1
Predicted Positive	C(1,0)	1	C(1,1)	0

Cost Matrix

C(i,j) ➔ cost of assigning an observation of class j to class i

C(0,0) and C(1,1) ➔ Cost of correct classification, usually 0

C(0,1) and C(1,0) ➔ Cost of FN and FP, respectively, usually 1

	Real Negative	Real Positive
Predicted Negative	C(0,0) 0	C(0,1) 100
Predicted Positive	C(1,0) 1	C(1,1) 0

Given a cost matrix, an observation should be classified into the class that has the minimum cost

Conditional Risk – Expected Cost

- R(i | x) is the expected cost of classifying an observation into class i
- P(j | x) is the probability of an observation of being of class j
- C(i,j) is the cost of assigning an observation of class j to class i

$$R(i|x) = \sum_{j=1}^{M} P(j|x) \cdot C(i, j),$$

Conditional Risk – Expected Cost

- R(i | x) is the expected cost of classifying an observation into class i
- P(j | x) is the probability of an observation of being of class j
- C(i,j) is the cost of assigning an observation of class j to class I

$$R(i|x) = \sum_{j=1}^{M} P(j|x) \cdot C(i, j),$$

An observation should be classified into the class that has the minimum cost or minimum risk

Conditional Risk – Expected Cost

For observation 1, assuming binary classification, and 1 minority class

- P(0 | 1) = 0.8 and P(1 | 1) = 0.2
- C(1,0) = 1 and C(0,1) = 10

- R(0,1) = P(0, 1) * C(0,0) + P(1 | 1) * C(0,1)
- R(0,1) = 0.8 * 0 + 0.2 * 10
- R(0,1) = 2

- R(1,1) = P(0, 1) * C(1,0) + P(1 | 1) * C(1,1)
- R(1,1) = 0.2 * 1 + 0.8 * 0
- R(1,1) = 0.2

$$R(i|x) = \sum_{j=1}^{M} P(j|x) \cdot C(i, j),$$

Conditional Risk – Expected Cost

For observation 1, assuming binary classification, and 1 minority class

- P(0 | 1) = 0.8 and P(1 | 1) = 0.2
- C(1,0) = 1 and C(0,1) = 10
- R(0 | 1) = 2 and R(1 | 1) = 0.2

$$R(i|x) = \sum_{j=1}^{M} P(j|x) \cdot C(i, j),$$

Even though P(0 | 1) > (1 | 1), R(0 | 1) > R (1 | 1) ➔ we predict class 1

Conditional Risk – Expected Cost

The classifier will classify an instance x into positive class if and only if:

$$R(1 | x) <= R(0 | x)$$

Conditional Risk – Expected Cost

The classifier will classify an instance x into positive class if and only if:

$$R(1 | x) <= R(0 | x)$$

- $R(0,1) = P(0, 1) * C(0,0) + P(1 | 1) * C(0,1)$
- $R(1,1) = P(0, 1) * C(1,0) + P(1 | 1) * C(1,1)$

$$P(0 | x)C(1,0) + P(1 | x)C(1,1) \leq P(0 | x)C(0,0) + P(1 | x)C(0,1)$$

Conditional Risk – Expected Cost

The classifier will classify an instance x into positive class if and only if:

$$R(1|x) <= R(0|x)$$

- R(0,1) = P(0, 1) * C(0,0) + P(1|1) * C(0,1)
- R(1,1) = P(0, 1) * C(1,0) + P(1|1) * C(1,1)

$$P(0|x)C(1,0) \leq P(1|x)C(0,1)$$

Conditional Risk – Expected Cost

$$P(0|x)C(1,0) \leq P(1|x)C(0,1)$$

$$(1-P(1|x))\ C(1,0) \leq P(1|x)C(0,1)$$

$$C(1,0)-P(1|x)\ C(1,0) \leq P(1|x)C(0,1)$$

$$C(1,0) \leq P(1|x)C(0,1) + P(1|x)\ C(1,0)$$

$$C(1,0) \leq P(1|x)(C(0,1) + C(1,0))$$

Conditional Risk – Expected Cost

$$P(0|x)C(1,0) \le P(1|x)C(0,1)$$

$$(1-P(1|x))\, C(1,0) \le P(1|x)C(0,1)$$

$$C(1,0) - P(1|x)\, C(1,0) \le P(1|x)C(0,1)$$

$$C(1,0) \le P(1|x)C(0,1) + P(1|x)\, C(1,0)$$

$$C(1,0) \le P(1|x)(C(0,1) + C(1,0))$$

Conditional Risk – Expected Cost

$$P(0|x)C(1,0) \le P(1|x)C(0,1)$$

$$(1-P(1|x))\, C(1,0) \le P(1|x)C(0,1)$$

$$C(1,0) - P(1|x)\, C(1,0) \le P(1|x)C(0,1)$$

$$C(1,0) \le P(1|x)C(0,1) + P(1|x)\, C(1,0)$$

$$C(1,0) \le P(1|x)(C(0,1) + C(1,0))$$

Conditional Risk – Expected Cost

$$P(0|x)C(1,0) \leq P(1|x)C(0,1)$$

$$(1-P(1|x))\ C(1,0) \leq P(1|x)C(0,1)$$

$$C(1,0) - P(1|x)\ C(1,0) \leq P(1|x)C(0,1)$$

$$C(1,0) \leq P(1|x)C(0,1) + P(1|x)\ C(1,0)$$

$$C(1,0) \leq P(1|x)(C(0,1) + C(1,0))$$

Conditional Risk – Expected Cost

We can transform the risk back to a probability, so now, I can determine the threshold of probability above which I can "confidently" classify an observation as a member of class 1

$$\frac{C(1,0)}{(C(0,1) + C(1,0))} \leq P(1|x)$$

Conditional Risk – Expected Cost

- $C(1,0) = 1$ and $C(0,1) = 10$

$$\frac{C(1,0)}{(C(0,1) + C(1,0))} \leq P(1|x)$$

$$\frac{1}{(10 + 1)} \leq P(1|x)$$

THANK YOU

www.trainindata.com