

Lista de Exercícios 1 - AM2 - Aulas 01, 02 e 03

Assuntos: Combinação de classificadores, Classificação multirrótulo e hierárquica.

Professor: Murilo Naldi

- 1) Qual dos algoritmos seguintes não é / são exemplos de combinação de classificadores?
 1. AdaBoost
 2. SVM
 3. Random Forest
 4. Gradient Boosting
- 2) O que é necessário para que uma combinação de modelos de classificação tenha sucesso? Existe garantia de sucesso?
- 3) Qual das afirmações é / são falsas sobre a combinação de classificadores?
 1. A combinação de classificadores se beneficia quando os modelos individuais cometem erros independentes.
 2. A combinação de classificadores heterôgeneos pode apresentar vantagens como a redução de sobreajuste (*overfitting*) e aumento do desempenho preditivo.
 3. O algoritmo *Random Forest* é um exemplo de combinação de classificadores onde cada modelo tenta superar os erros do modelo anterior, atribuindo um peso maior às classificações incorretas sempre que um novo modelo é ajustado.
 4. O algoritmo *Random Forest* é um exemplo de combinação de classificadores que utiliza a técnica Bagging.
- 4) Como combinar previsões de modelos diferentes? Dê exemplos.
- 5) Qual das alternativas seguintes está / estão corretas em relação aos benefícios da combinação de classificadores?
 1. Melhor capacidade de generalização.
 2. Melhor interpretabilidade dos resultados.
 3. Menor tempo de inferência.
- 6) Quais são as melhores formas de se gerar modelos distintos, especialmente quando a combinação é homogênea?
- 7) Qual dos seguintes é / são verdadeiros sobre os classificadores fracos usados em comitês de classificadores?
 1. Possuem baixa variância e não costumam sofrer com sobreajuste (*overfitting*).
 2. Possuem alto viés, o que causa subajuste (*underfitting*) em problemas mais complexos.
 3. Possuem alta variância e não costumam sofrer com sobreajuste (*overfitting*).
- 8) A maioria dos problemas de classificação associa uma única classe a cada exemplo ou instância. No entanto, existem muitas tarefas de classificação em que cada instância pode ser associada a uma ou mais classes. Esse tipo de tarefas pertencem ao grupo conhecido como:
 1. Classificação multiclasse
 2. Classificação multirrótulo

3. Classificação multivariada
4. Classificação multisaída

9) One-vs-all e Label Powerset são, respectivamente, exemplos de técnicas para classificação:

1. Multirrótulo e multisaída
2. Multirrótulo e multiclasse
3. Hierárquica e multirrótulo
4. Multiclasse e multirrótulo

10) Explique as vantagens e desvantagens das técnicas de Relevância Binárias e Label Powerset. Em quais casos são vantajosas e quais não são?

11) Qual dos seguintes não é / são técnicas de classificação multirrótulo:

1. Label Poweset
2. Relevância binária
3. ECOC
4. Cadeia de classificadores

12) Qual das afirmações abaixo é / são verdadeiras?

1. Após utilizar a estratégia *Label Powerset* para transformar nossos dados em uma tarefa de classificação multirrótulo, poderíamos utilizar o método One-vs-all para treinar classificadores.
2. A distância de *Hamming* é uma métrica de avaliação para tarefas de classificação multiclasse.
3. Na classificação multiclasse um único exemplo recebe exatamente um rótulo de um grupo de muitas classes possíveis.
4. A regressão multiobjetivo se assemelha a classificação multirrótulo em um contexto de problemas de regressão.

13) Treinar múltiplos modelos, combinar suas saídas (*scores*) e usá-las como um conjunto de treinamento para treinar um classificador a partir delas, é um exemplo de qual método de combinação de classificadores?

1. Encadeamento de classificadores
2. Sieriação
3. Generalização por pilha
4. Label Powerset

14) Qual das afirmações abaixo é / são verdadeiras?

1. O encadeamento de classificadores e a generalização por pilha são métodos de combinação de classificadores.
2. A relevância binária é um método que considera a dependência entre classes.
3. *ML kNN* e *RAKEL* são exemplos de algoritmos de aprendizado de máquina adaptados para problemas de classificação multirrótulo.
4. A Hamming loss é uma métrica para avaliação de classificadores multirrótulo.

Quanto maior seu valor, menos erros estão sendo cometidos pelo classificador.

15) Explique a diferença entre classificação multi-classe, multi-rótulo e hierárquica. São conceitos excludentes? Justifique.

16) Qual das afirmações é / são falsas sobre a classificação hierárquica?

4. A classificação hierárquica pode ser realizada somente de maneira completa, realizando a predição obrigatoriamente até os nós folha.
5. A classificação hierárquica pode ser multirrótulo caso haja mais de um caminho possível na hierarquia.
6. Árvore e Grafo acíclico direto são exemplos de estruturas de hierarquia utilizadas na classificação hierárquica.

17) Qual das afirmações é / são falsas sobre a classificação hierárquica?

5. A abordagem de classificadores locais possui um problema relacionado a propagação de erros de níveis superiores para o restante da hierarquia.
6. O método Global (ou Big-Bang) normalmente necessita de menor poder computacional em relação as outras abordagens.
7. A classificação plana é uma estratégia que ignora completamente a hierarquia de classes.

18) Compare as abordagens local e global para uma estrutura de um classificador hierárquico. Quais são as vantagens e desvantagens de cada abordagem?