

Universidade Federal Rural de Pernambuco

Disciplina: Inteligência Artificial Turma: 2018.2

Prof: Valmir Macario Filho

Projeto para avaliação da disciplina de IA.

1ª FASE (2.0 pontos) – Criação do vetor de características:

Primeiro, escolha um dos temas de projetos descritos abaixo. Depois, pense no objetivo do problema que você escolheu.

1. Escolha um dos objetivos finais de classificar, agrupar ou otimizar.
2. Quais são as classes ou categorias do seu problema?
3. Quais são as características desse projeto que descrevem o objeto ou exemplo do problema para que esse mesmo objeto seja classificado como pertencente a uma determinada classe ou grupo. Se algoritmo genético, quais serão os atributos de seu cromossomo.
4. O que fará diferença para que objetos de classes ou grupos diferentes sejam diferenciados?
5. Se algoritmo genético ou de busca, quais serão as operações do seu algoritmo de otimização.

Faça um relatório respondendo à essas perguntas usando o formato de artigos da SBC. Comece com uma introdução, descrevendo seu problema em detalhes. Depois, escreva porque está resolvendo este problema, qual a motivação do problema? Depois, descreva como você pretende resolver esse problema respondendo às perguntas anteriores. Qual algoritmo pretende utilizar pra resolver seu problema? Quais os parâmetros você vai modificar pra melhorar os resultados? Crie um arquivo de texto no formato “arff” do weka para servir de entrada para o seu algoritmo.

DATA DE ENTREGA: 10/02/2018

2ª FASE (3.0 pontos) – Implementação do algoritmo de inteligência computacional

Após criar o vetor de características, temos que pensar qual algoritmo de inteligência computacional será utilizado para realizar a classificação ou agrupamento dos exemplos do nosso problema. Opções:

- | | |
|-------------------------|--------------------------|
| - KNN Ponderado | -Árvore de Bayes |
| - K-Means | -Árvore de Decisão |
| - K-Medoid | -Perceptron |
| -Clustering Hierárquico | -Perceptron Multicamadas |
| -Naive Bayes | -Algoritmos Genéticos |
| -Simulated annealing | -A* |

- Busca Tabu

-Outro de sua preferência acordado com o professor

Implemente um desses algoritmos para resolver seu problema. Complemente seu relatório descrevendo o algoritmo que você escolheu e inclua o pseudo-código do algoritmo.

DATA DE ENTREGA: 20/12/2018

3ª FASE (5.0 pontos) – Avaliação dos algoritmos

Após implementar seu algoritmo, avalie seu algoritmo utilizando pelo menos 3 parâmetros diferentes. A avaliação dos algoritmos será pela taxa de acerto no conjunto de teste. Complemente seu relatório com uma seção explicando detalhadamente quais foram os experimentos realizados. Em outra seção coloque tabelas ou gráficos com os resultados obtidos e faça uma análise desses resultados. Finalize com uma seção de conclusão. Prepare um seminário de 10 minutos apresentando o trabalho realizado, com enfoque no problema, no algoritmo, nos experimentos, resultados e conclusões.

DATA DE ENTREGA e SEMINÁRIOS: 10/01/2019

PROJETOS:

1. Reconhecimento de Caracteres.

Descobrir a letra que está na imagem. A base de dados MNIST possui uma base de dados com caracteres que representam números de 0 a 9.

<http://www.iro.umontreal.ca/~lisa/twiki/bin/view.cgi/Public/DeepVsShallowComparisonICML2007>

<http://yann.lecun.com/exdb/mnist/>

2. Classificação de Imagens:

São imagens de 10 objetos diferentes com resolução de 32x32 bits. As imagens podem ser encontradas em:

<http://www.cs.toronto.edu/~kriz/cifar.html>

3. Classificação de Texto

Classificação de notícias retiradas do site da Reuters.

<http://www.daviddlewis.com/resources/testcollections/reuters21578/>

4. Recomendação de Filmes

Base de dados com recomendação de filmes com várias informações sobre a popularidade de um filme.

<https://www.kaggle.com/tmdb/tmdb-movie-metadata>

5. <https://www.kaggle.com/tmdb/tmdb-movie-metadata> Notícias da BBC

Base de dados formada por 5 categorias de notícias do site BBC. O objetivo é identificar qual categoria (business, entertainment, politics, sport, tech) de cada notícia.

<http://mlg.ucd.ie/datasets/bbc.html>

6. Base de dados de análise de sentimentos sobre filmes. São 1000 exemplos de comentários negativos e 1000 exemplos de comentários positivos sobre filmes.
<https://www.cs.cornell.edu/people/pabo/movie-review-data/>
7. Base de dados de análise de sentimentos de tweets sobre produtos, personalidades entre outros. O objetivo é classificar o tweet entre 3 classes (negativo, neutro e positivo).
<http://help.sentiment140.com/for-students/>
8. Dados Abertos. Bases de dados com informações sobre a cidade de Recife, estados brasileiros e Brasil.
<http://dados.recife.pe.gov.br/group>
<http://www.dadosabertos.pe.gov.br/PortalDadosAbertos/public/pages/telaInicial.jsf>
<http://dados.gov.br/dataset>
9. ROBOCODE. Desenvolver um robô capaz de ganhar de outros numa batalha de robôs tanques utilizando algoritmos de Busca e\ou genéticos.
<http://robocode.sourceforge.net/>
10. Sites com mais opções de projeto:
<https://www.kaggle.com/>
<https://toolbox.google.com/datasetsearch>
<https://console.cloud.google.com/marketplace/browse?q=public%20data&filter=solution-type:dataset&pli=1>