

Sistemas Inteligentes Aplicados

Trabalho: Biblioteca scikit-learn (sk-learn)

Parte 1 – Execute os treinamentos (peso 5)

1. Rodar o seguintes 10 classificadores abaixo usando a biblioteca scikit-learn, para o dataset abaixo descrito. Sugestão: usar jupyter notebooks do Google Colab.

1.1 Árvore de Decisão CART (padrão).

1.2 Support Vector Machine (padrão).

1.3 Support Vector Machine (mudando o kernel por outro).

1.4 Knn (padrão)

1.5 Knn (mudando o valor de k)

1.6 Knn (mudando a medida de distância)

1.7 Gaussian Naive Bayes (padrão). Obs: para atributos quantitativos, pode remover atributos qualitativos para simplificar o problema.

1.8 Categorical Naive Bayes (padrão). Obs: para atributos qualitativos, pode remover atributos quantitativos para simplificar o problema.

1.9 Regressão Linear (quadrados mínimos). Obs: escolher dataset para regressão.

1.10 Perceptron.

2. Conforme o algoritmo, é possível que atributos qualitativos ou quantitativos não sejam aceitos. Você pode remover atributos incompatíveis com o algoritmo em mãos.

3. Você pode usar o weka para exportar os datasets para o formato CSV.

4. Você pode usar o pandas para importar os datasets ou fazer um código próprio para isso.

5. Envie o código e os datasets utilizados.

Parte 2 – Elabore o relatório (peso 5)

Faça um relatório com no mínimo 500 palavras (usar “contar palavras” do Word ou similar) contendo:

1. Uma breve descrição do dataset.

2. Uma descrição sobre o scikit-learn.

3. Os resultados obtidos para cada teste realizado
4. O que fazem os parâmetros alterado para nos itens 1.2, 1.4, e quais valores foram usados.
5. Tabelas, figuras, legendas de tabelas e figura, capa, contra-capas, fórmulas, pseudo-código, etc não contam como palavras.
6. Não há formato obrigatório, se estiver em dúvida sobre qual formato usar, use report (no latex) ou abnt (latex, word, etc)
7. Enviar .tex ou .docx ou equivalente para contagem de palavras, não apenas o .pdf.

Avaliação

1. O trabalho é individual.
2. Detecção de plágio, mesmo que em pequenas porções de texto ou slides **anularão o trabalho**. Se estiver em dúvida sobre o que caracteriza plágio, consulte o professor.

DataSet

Um acidente vascular cerebral (AVC) é uma condição médica na qual o fluxo sanguíneo deficiente para o cérebro causa a morte celular. Existem dois tipos principais de acidente vascular cerebral: isquêmico, devido à falta de fluxo sanguíneo, e hemorrágico, devido ao sangramento. Ambos fazem com que partes do cérebro parem de funcionar corretamente. Os sinais e sintomas de um acidente vascular cerebral podem incluir incapacidade de se mover ou sentir em um lado do corpo, problemas de compreensão ou fala, tontura ou perda de visão de um lado. Os sinais e sintomas geralmente aparecem logo após o acidente vascular cerebral. Se os sintomas durarem menos de uma ou duas horas, o AVC é um ataque isquêmico transitório (AIT), também chamado de miniderrame. Um acidente vascular cerebral hemorrágico também pode estar associado a uma forte dor de cabeça. Os sintomas de um acidente vascular cerebral podem ser permanentes. As complicações a longo prazo podem incluir pneumonia e perda do controle da bexiga.

O principal fator de risco para acidente vascular cerebral é a pressão alta. Outros fatores de risco incluem colesterol elevado no sangue, tabagismo, obesidade, diabetes mellitus, um AIT anterior, doença renal terminal e fibrilação atrial. Um acidente vascular cerebral isquêmico é tipicamente causado pelo bloqueio de um vaso sanguíneo, embora também existam causas menos comuns. Um acidente vascular cerebral hemorrágico é causado por sangramento diretamente no cérebro ou no espaço entre as membranas do cérebro. O sangramento pode ocorrer devido a um aneurisma cerebral rompido. O diagnóstico geralmente é baseado em um exame físico e é apoiado por imagens médicas, como tomografia computadorizada ou ressonância magnética. Uma tomografia computadorizada pode descartar sangramento, mas não necessariamente descarta isquemia, que no início normalmente não aparece em uma tomografia computadorizada. Outros exames, como eletrocardiograma (ECG) e exames de sangue, são feitos para determinar os fatores de risco e descartar outras possíveis causas. Baixo nível de açúcar no sangue pode causar sintomas semelhantes.

A prevenção inclui a diminuição dos fatores de risco, cirurgia para abrir as artérias do cérebro em pessoas com estreitamento carotídeo problemático e varfarina em pessoas com fibrilação atrial. Aspirina ou estatinas podem ser recomendadas pelos médicos para prevenção. Um acidente vascular cerebral ou TIA geralmente requer cuidados de emergência. Um acidente vascular cerebral isquêmico, se detectado dentro de três a quatro horas e meia, pode ser tratável com um medicamento que pode quebrar o coágulo. Alguns derrames hemorrágicos se beneficiam da cirurgia. O tratamento para tentar recuperar a função perdida é chamado de reabilitação de acidente vascular cerebral e, idealmente, ocorre em uma unidade de acidente vascular cerebral; no entanto, estes não estão disponíveis em grande parte do mundo.

Informações de atributo

- 1) sexo: "Masculino", "Feminino" ou "Outro"
- 2) idade: idade do paciente
- 3) hipertensão: 0 se o paciente não tem hipertensão, 1 se o paciente tem hipertensão
- 4) doença cardíaca: 0 se o paciente não tiver nenhuma doença cardíaca, 1 se o paciente tiver uma doença cardíaca
- 5) casou-se: "Não" ou "Sim"
- 6) tipo de trabalho: "crianças", "Funcionário público", "nunca trabalhou", "particular" ou "autônomo"
- 7) tipo de residência: "rural" ou "urbano"
- 8) avgglucoselevel: nível médio de glicose no sangue
- 9) IMC: índice de massa corporal
- 10) smoking_status: "ex-fumante", "nunca fumou", "fuma" ou "Desconhecido"*
- 11) acidente vascular cerebral: 1 se o paciente teve um acidente vascular cerebral ou 0 se não

*Observação: "Desconhecido" em smoking_status significa que a informação não está disponível para este paciente

Fontes de dados reais.

Observe: os dados estão em escalas diferentes.