Aula dia 04/08/2023

* Muito difícil definir IA, pois cada autor tem a sua, porém existem algumas compressões comuns
* Linha do tempo

  Descrição gerada automaticamente com confiança baixaQuatro quadrantes do Russel:

OBS:

* + Pensar e agir como humanos são os pontos mais difíceis, porém com mais avanços recentemente
* **Pensar racionalmente:**
  + IA clássica/ simbólica
  + Preposições lógicas
  + Qualquer problema descrito através de lógica pode ser resolvido
  + Eliza – robô de comunicação
  + Assistentes virtuais
* **Pensar humanamente:**
  + É preciso entender como os humanos pensam
  + Redes neurais
* **Agir humanamente:**
  + Processamento de linguagem natural
  + Representação do conhecimento
  + Raciocínio
  + Teste de turing
* **Agir racionalmente:**
  + Um agente percebe seu ambiente e age através de atuadores
* Interface gráfica do usuário, Texto, Aplicativo, Email

  Descrição gerada automaticamenteResumos de definições:

Interface gráfica do usuário, Texto, Aplicativo, Email

Descrição gerada automaticamente

Aprendizado de máquina

* Área da IA que trabalha com a experiencia – qualquer sistema que aprende com exemplos

Aula do dia 08/08/2023 – Aprendizado de máquina

* Definições:
  + Construção de programas de computador que melhoram seu desempenho por meio de experiência
  + Um programa aprende a partir da experiência E, em relação a uma classe de tarefas T, com medida de desempenho P, se seu desempenho em T, medido por P, melhora com E (Tom.M. Mitchell, 1997)
  + O aprendizado de máquina é o treinamento de um modelo a partir de dados que generalizam uma decisão contra uma medida de desempenho (Jason Brownlee)
* Representatividade nas bases de dados:

Para termos variações nos atributos, devemos ter um dado de cada um, assim com mais atrbibutos devemos ter a combinação deles

Tabela

Descrição gerada automaticamenteEX: Sol + Quente + Alta + Não + Sim ...

Se temos 300 atributos e apenas 118 instancias podemos dizer que esses dados não são representativos

* Atributos:
  + Nominais: são aqueles usados para se dar um rótulo para as variáveis sem valores quantitativos. Exemplos comuns são masculino/feminino
  + Numéricos: Idade, peso, tamanho, binário
* Tipos de problemas:
  + **Classificação**:
    - Mais utilizados
    - Queremos prever ou descrever uma classe
    - Atributo nominal da previsão
    - Já teremos um rótulo pré-definido
  + **Regressão**:
    - Semelhante ao problema anterior
    - Atributo numérico da previsão
  + **Agrupamento(clusterização)** 
    - Agrupar as instancias de acordo com os atributos de entrada
    - Não é conhecido o atributo de classificação
    - Usado para identificação de perfis
  + **Regras de associação:**
    - Busca semelhante/associações entre os elementos
* Tipos de aprendizado:
  + **Supervisionado**:
    - Indutor recebe conjunto de exemplos na forma entrada e rotulo
    - Técnicas: Redes neurais , Maquinas de vetores suporte, Arvore de decisão
  + **Não-supervisionado**:
    - Apenas atributos de entrada
    - Encontrar agrupamentos (todos de agrupamento)
    - Técnicas: Rede neurais do tipo mapas auto-organizaveis e K-means
  + **Semi-supervisionado:**
    - Algumas instancias com rótulos e outras não
  + **Reforço**:
    - Maximização das recompensas do resultado
    - Recompensas e penalizações
  + **Deep Learning**:
    - Rede neural com muitas camadas. Para cada reconhecendo situações com ordem de complexidades maiores
* KDD
* Diagrama

  Descrição gerada automaticamenteMineração de dados é encontrar padrões nos dados – com machine learning ou outras técnicas

Aula dia 11/08/2023 – Árvore de decisão

Introdução:

* Gerada por meio de aprendizado
* Base fornecida com instancias e atributos de entrada, e a árvore irá classificar no atributo de classificação.
* Linha do tempo

  Descrição gerada automaticamenteEX(esquerda = TRUE e direita = FALSE):
* Podemos observar que houve um descarte de atributos, aém de um erro de 5 instancias
* O número de regras são é aquantidade de folhas
* Quanto o atributo é numérico, a arvore tende a se dividir em dois, mas isso não é uma regra

Slide 01:

* Pode ser árvore de **classificação** (ID3 C45 CART) ou **regressão** (CART)
* Simples fácil e interpretada
* Complexidade: numero total de nodos, folhas , profundidade da arvore e numero de atributos
* Disjunção de conjunções – ORS e ANDS
* Hierarquicamente arranjadas
* Uma imagem contendo Diagrama

  Descrição gerada automaticamenteEstrutura:
* Qualquer função booleana pode ser escrita como uma arvore de decisao

Aula dia 18/08/2023

* Como gerar arvores:
  + Sempre escolhe nós que serão mais simples
  + Como calcular raiz?
    - ID3 – Ganho(atributo) = entropia(classe) – entropia(atributo)

Tela de celular com texto preto sobre fundo branco

Descrição gerada automaticamente com confiança médiaAula dia 22/08/2023

* Para gerarmos o segundo nível da arvore:
  + Desconsiderar o restante das pessoas, e calcular tudo de novo
* Particularidades:
  + Tentativa de encontra menor ramificações possíveis, assim quando temos um empate de ganho devemos escolher o com menor quantidade de opções, depois aleatório
  + A base do log é a quantidade de resposta da classe(Sim,Não,Talvez) – seria log base 3
* Gráfico

  Descrição gerada automaticamenteOs termos de avaliação vêm da matriz de confusão – todo problema de classificação tem isso:
* Termos de avaliação do modelo – Sempre separamos uma % da base para teste:
  + **Recall(sensibilidade/revocação):** Quantidade de acertos da classe – VP / VP + FN 🡪 FN mais importante
  + **Precisão:** Quão preciso ele é VP/VP+FP 🡪 FP mais importante
  + **F1 Score:** Media harmônica das duas anteriores – 2 x recall x precisao / precisao+recall
  + **Acurácia =** VP + VN / VP+VN+FN+FP

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **recall** | **precisão** | **F1 score** | **TVP** | **TFN** | **TFP** | **TVN** |
| **setosa** | 20/20 - 1 | 20/20 - 1 | 1 | 1 | 0 | 0/51 | 51/51 |
| **Versicolor** | 21/26 0,80 | 21/22 – 0,95 | 0,86 | 0,8 | 0,2 | 1/50 | 49/50 |
| **virginica** | 24/25-0,96 | 24/29 – 0,82 | 0,88 | 0,96 | 0,04 | 5/46 | 42/46 |

* **Exercício:**

**Aula dia 25/08/2023**

Tabela

Descrição gerada automaticamente

* Existem alguns casos que devemos analisar o que é pior falso negativo e falso positovo – lembrar exemplo da aids
* Fazendo árvore em python:
  + Biblioteca – sklearn 🡪 faz o CART , nenhum outro
  + Transformação de dados:
    - Dado binário – transforma para 0 ou 1 – LabelEncoder faz isso
    - Dado nominal ordinal: transforma em 0 1 2 3 4 - LabelEncoder faz isso
    - Dado nominal não ordinal - OneHotEncoder:
      * Tabela

        Descrição gerada automaticamenteBinárizar o atributo - se tiver muita opção será diferente
      * LabelEnconder usará ordem alfabética

Aula dia 29/08/2023

* Texto

  Descrição gerada automaticamenteCobertura de regras:
  + Global: 4/9
  + Por classe: 4/5
  + Analisar a qualidade da regra – quanto maior melhor a regra
* 
  + Random\_state 🡪 garante que a arvore será sempre a mesma

Aula dia 05/09/2023

* Quando o atributo é numérico binário 🡪 faz uma discretizarão e uma conta só
* Quando o atributo é numérico sem ser binário 🡪 Calcula média entre dois primeiros e faz corte do resto, depois tenta outros cortes com a finalidade de aumentar o ganho
* Aprendizado em conjunto / Ensembles learning – Random Forest
  + Resultado bons parecidos com rede neural
  + Métodos que geram muitos classificadores e combinam os seus resultados
  + É amplamente aceito que o desempenho de um conjunto de muitos classificadores fracos é geralmente melhor do que um único classificador, dada a mesma quantidade de informação de treinamento
* Random Forest:
  + Mesma base de dados e várias arvores diferentes – dataset diferente
  + Diagrama

    Descrição gerada automaticamenteCombinação de predições:
  + Tecnica de amostragem bootstrap:
    - Amostragem aleatória com reposição para gerar diversidade
  + Texto

    Descrição gerada automaticamenteSimples, fácil de gerar e rápido
  + Não interpretado 🡪 principal desvantagem
  + Aleatório:
    - Seleção das instâncias – reposição
    - Quantidade de atributos

08/09/2023

* Boosting:
  + Caso incorreto aumentamos o peso
  + Vai gerando arvores novas com pesos para instancias incorretas

Estudo Individual 18/09/2023 – Naive Bayes

* Abordagens probabilística
* Interface gráfica do usuário, Texto

  Descrição gerada automaticamenteUma imagem contendo Gráfico

  Descrição gerada automaticamenteTabela de probabilidade:

Balanceamento de dados:

* Definição de desbalanceamento de dados: dados de um subconjunto  
  das classes aparecem com uma frequência maior que os  
  dados das demais classes.
* Técnicas:
  + Redefinir o tamanho do conjunto de dados
    - Neste caso, podemos tanto adicionar instância à classe minoritária (métodos oversampling) quanto remover instâncias da classe majoritária (métodos undersampling)
  + Utilizar diferentes custos de classificação para as diferentes classes
  + Induzir um modelo para uma classe

22/09/2023 – Regras de associação

* Regras de associação com regras de classificação
* Podem ser encontradas da mesma forma que regras de classificação
* Associar qualquer coisa com qualquer coisa
* Mineração de regras de associação:
  + O grau de incerteza é dado pela confiança da regra
* Apriori:
  + Confiança mínima e suporte mínimo
  + Suporte 🡪 frequência
  + Texto, Carta

    Descrição gerada automaticamenteConfiança da regra:
  + Passo a passo:
    - Tabela

      Descrição gerada automaticamente1) Calcular suporte de tudo:

17/10/2023 – Agrupamento (Algoritmo K means)

* Normalização de dados: colocar todos os dados em uma mesma faixa
* Pré-processamento:
  + Normalizar dados
  + Escolher uma proximidade que não depende de da magnitude dos dados
  + Eliminar outliers – primeira coisa a fazer
  + Seleção de atributos
* Sempre precisamos validar o agrupamento
* Usamos muitas vezes agrupamento para fazer o balanceamento

27/10/2023 – BackPropagation rede neural

* Para resolver problemas não linearmente separáveis.
* Apresentam uma ou mais camadas intermediarias de neurônios e uma camada de saída 🡪 (número de saídas = colunas) recomendado
* Backpropagation:
  + Algoritimo de redes multi-camadas mais difundido.
  + Aprendizado supervisionado por correção de erros:

07/11/2023 – Métodos de busca

* Tipos:
  + Cega(não analisa o problema) – g(n):
    - Profundidade
    - Largura
    - Dijkstra
    - OBS: Custo uniforme
  + Heurística – h(n):
    - Gulosa
    - A\*
* Busca gulosa:
  + Heurística é determinada analisando o problema
  + Valor da heurística não é somada
  + Não garante a solução ótima , ou seja , o menor caminho
  + Muito rápida
* Distância de Manhattan:
  + Heurística mais usada
  + a menor distância possível que um carro é capaz de percorrer numa malha urbana reticulada ortogonal

Aula dia 14/11/2023 - minimax

* Consideraremos jogos com dois jogadores: – MAX e MIN – MAX faz o primeiro movimento e depois eles se revezam até o jogo terminar.
* Um jogo pode ser definido como um problema de busca com: – estado inicial – função sucessor (-> movimento, estado) – teste de término – função utilidade: dá um valor numérico para os estados terminais
* Teoria: gerar a arvore inteira até o final , já não pratica usamos até certa altura
* Diagrama, Esquemático

  Descrição gerada automaticamenteMAX – MIN
* Max sempre será o maior valor e o min o menor
* Triangulo pra cima = max e triangulo pra baixo = min
* Problema: estimativa tem que ser boa e sempre tem que ser a melhor jogada para funcionar + eficiência ruim pois entra em vários nós atoa
* Interface gráfica do usuário, Texto, Aplicativo

  Descrição gerada automaticamentePoda Alpha-Beta:

21/11/2023 – redes neurais profundas e convulsionais

* Uma rede neural articifial que possui muitas camadas é uma deep neural network
  + processam esses dados por meio de muitas camadas de transformações não lineares dos dados de entrada para calcular uma saída de destino
* Gráfico, Diagrama

  Descrição gerada automaticamenteDeep learning é uma continuidade da rede neural
  + Rede neural convulsionais – anos 70
  + Antigamente não podíamos dar input de imagem, então tínhamos que extrair os dados da imagem , pre processar e depois jogar no modelo
* Redes neurais convolucionais:
  + algoritmo de Aprendizado Profundo que pode captar uma imagem de entrada, atribuir importância (pesos e vieses que podem ser aprendidos) a vários aspectos / objetos da imagem e ser capaz de diferenciar um do outro.
  + rede neural convolucional (CNN/ConvNet) é uma classe de redes neurais profundas, mais comumente aplicada para analisar imagens visuais.
  + Etapas:
    - 1- Operador de convolução:
      * Passar por vários kernel – que são extratores de características
      * Com o mapa de características (filter map), a imagem fica menor para facilitar o processamento
      * Alguma informação sobre a imagem pode ser perdida, porém o propósito é detectar as partes principais (quanto maior os valores melhores)
      * Interface gráfica do usuário, Aplicativo

        Descrição gerada automaticamenteO mapa de características preserva as características principais da imagem (olho, boca, etc)
      * Shift de 1 e faz a mesma coisa
      * Passa por uma função de ativação(**RELU**) 🡪 máximo (0 e z) – negativo = 0
    - 2 – Pooling
      * Seleciona as características mais relevantes da imagem (reduz overfitting e ruídos)
      * Interface gráfica do usuário

        Descrição gerada automaticamente com confiança médiaPode-se usar o máximo, o mínimo a média dos valores, dentre outros
      * Mais usual é o max pulling
      * Shift de 2
    - 3 – Flattening:
      * Gráfico, Diagrama

        Descrição gerada automaticamenteColoca a matrix de forma linear e joga na rede neural densa

24/11/2023 – mineração de texto

* Dado não estruturado

Diagrama

Descrição gerada automaticamente28/11/2023

Diagrama

Descrição gerada automaticamente

* Extração de termos:
  + Análise(parsing)
  + Remoção de termos comuns 🡪 artigos, preposições , conjunções
  + Análise sintática🡪 sinônimos , polissemia
  + Redução morfológica 🡪 lematização e stemização
* Calculo dos atributos:
  + Foto em preto e branco

    Descrição gerada automaticamente com confiança média
  + Como calcular frequência:
  + Texto, Tabela

    Descrição gerada automaticamente com confiança médiaResultados:
    - Os textos são representados pelos valores TFxIDF de cada
    - termo.
    - TFxIDF igual a zero indica termo não relevante.
    - TFxIDF maior indica termo mais relevante.
    - Textos podem ser agrupados e categorizados com base no vetor de valores TFxIDF
  + Interface gráfica do usuário, Texto, Aplicativo

    Descrição gerada automaticamenteRedução de dimensionalidade: