coggle
made for free at coggle.it

Introdução a Inteligência Artificial

Introdução, Diferenças e Conceitos Básicos

Tipos de IA

Inspirada em Humanos

- Cognitiva e Emocional
- Além de considerar questões cognitivas também consideram as emoções para tomadas de decisões
- Exemplos: Chatbots; Automação de Recrutamento

Humanizada

- Autoconsciente e adaptável em interações
- Falharia Criatividade para ser um humano
- Cognitiva, Emocional e Social

Analítica

- Somente Cognitiva
- Usa experiência passada para tomar decisões futuras
- Exemplos: Sistema Antifraude Reconhecimento Imagem Carros Autômatos

Divisões da IA

IA Fraca

1. Automação Inteligente
   - Tarefas Repetitivas
   - Requerem Inteligência
   - Ex. Sistema Automático de Separação de Recicláveis
2. Assistente Inteligente
   - Revisam e Identificam padrões em dados históricos
   - Ajudam pessoas executar tarefas mais rapidamente
   - Ex. Deep Learning, Processador de linguagem natural

IA Forte

3. Inteligência Aumentada
   - Ajudam prever futuro incerto
   - Ex. Sistema para simular cenários pelas políticas climáticas
   - Consegue perceber outros agentes e entidades e consideram isso para ajustar seu comportamento
4. Inteligência Autônoma
   - Sistemas de tomada de decisão sem Intervenção humana
   - Ex. Sistemas que identificam padrões de consumo e custo de aquecedor para adaptar o uso
   - Possui autoconsciência

Teste de Turing

Visualização de Dados

- Histograma
  - Usado para resumir a distribuição de uma amostra de dados.
  - Essencialmente um gráfico de barras, onde cada categoria no eixo x representa um intervalo de valores de observação
- Gráfico de Linhas
  - Úteis para apresentar dados de séries temporais, bem como quaisquer dados de sequência, onde há uma ordenação entre observações.
  - Usado para apresentar observações coletadas em intervalos regulares.
- Plot de Barras
  - Usado para apresentar quantidades relativas para várias categorias.
  - São úteis para comparar várias grandezas pontuais ou estimativas.
- Gráfico de Dispersão
  - Usado para resumir a relação entre duas amostras de dados emparelhadas.
  - Os eixos são distribuídos entre as amostras e os pontos são cada observação

Datasets, Tipos de Dados, Atributos e Features

Aprendizado

Otimização

- Ajustar Parâmetros Para
  - Ou Diminuir Erro
  - Ou Aumentar Recompensa

Representação

- Overfitting (Muito Complexo)
- Bem Ajustado
- Underfitting (Muito Simples)

Avaliação

- Separar Hipóteses - Boas das Ruins
- Medir Performance - Distancia das Amostras com a Função

Informação

- Não é conhecimento
- São dados com alguma estruturação ou organização
- Atribui significado a realidade mediante seus códigos e o conjunto de dados
- São dotados de relevância e propósito
- Exige consenso em relação ao significado
- Ex. Através dos dados de exemplo a informação é que está mais frio que o normal

Dados

- Problemas Com Dados
  - Omissos
    - Não foram salvos ou não podem ser lidos
    - Valores nulos
  - Ruidosos
    - Valores armazenados pouco confiáveis, aleatórios ou improváveis
    - Idade: 180 anos
      CPF: 999.999.999-99
  - Codificação
    - Dados armazenados em codificações diferentes
    - Um registro em ASCII e outro em Base64
  - Conflitantes
    - Dados divergentes para uma mesma observação
    - País: EUA
      Cidade: São Paulo
- Requerem Interpretação para ser informação.
- Valores ou ocorrências em um estado bruto
- O Conjunto de Dados pode constituir uma Informação
- São registros soltos, aleatórios, sem quaisquer análise
- Isoladamente não podem transmitir uma mensagem ou representar algum conhecimento
- Ex. A temperatura atual é 15ºC em Belo Horizonte às 12:00 do dia 12 de Janeiro
- Tipos de Dados
  - Categóricos ou Qualitativos (Categorias podem ser expressadas em forma de números)
    - Nominais
    - Ordinais
  - Numéricos ou Quantitativos
    - Discretos
    - Continuos
      - Intervalos
      - Proporções

Metadados

- Ex: Nome, Artista e Álbum acerca de uma música
- São dados sobre outros dados
- Pode facilitar o entendimento de um dado ou sua organização

Atributos

- O conjunto de atributos representa uma Entidade
- Os atributos são considerados metadados
- Tipos
  - Objeto (nome -> valor)
  - Arquivo (nome, extensão)
  - Banco de Dados (colunas + regras -> valores)

Conhecimento

- Através de um conjunto de informações gera uma ação ou o entendimento das causas de uma ação.
- Ex. Através do exemplo de informação com o conhecimento chega na decisão de será necessário sair com um casaco.
- É a informação processada e transformada em experiência pelo indivíduo.
- Através dele que determina-se quais são as decisões a tomar
- É gerado através de reflexão, síntese e contexto.

Features

- Propriedade mensurável de um fenômeno observado
- Atributo útil para solução de problema
- É crucial para algoritmo efetivo
- São numéricos
- Tipos de Features
  - Categóricas
    - Representam uma característica
    - Ex: Idioma ou Gênero
  - Textuais
    - São elementos de texto
    - Ex: Frases, Nomes ou Endereços
    - Transformações Comuns
      - Contagem de Palavras
      - Remoção de Acentos e Caracteres Especiais
      - Conversão em Upper/Lower Case
      - Remoção de pontuação
      - Remoção de Stopwords
  - Imagem / Audio / Vídeo
    - Depende de um processamento mais específico
    - Ex: Frequência de Áudio ou Textura de Imagem
  - Derivadas
    - Combinar, criar ou remover feature que descreve problema
    - Ex: Razão entre duas features (X/Y)

Engenharia de Features

- Transformar dados brutos em características que melhor representem o problema para que o algoritmo possa compreender
- Melhorar resultados de modelos de Maching Learning
- Objetivo é a Generalização
- Feature Selection
  - Remove features irrelevantes para o problema
  - Processo automática para selecionar subconjunto de features
  - Métodos de regularização como o LASSO e o RIDGE são considerados feature selection
  - Diminui Overfitting, Reduz Tempo de Treino e Aumenta a Precisão
- Feature Importance
  - Avaliar as features através de pontuações que podem ser atribuídas através de algoritmos
  - Auxilia na extração e construção de novas features similares
  - Garantir que use features que sejam realmente úteis para o modelo
  - Derivar features consideradas importantes para buscar novos contextos
- Feature Extraction
  - Geração automática de novas features a partir do dados bruto
  - Redução de Dimensionalidade
  - Diminuir número de recursos necessário para processamento sem perder informações
  - Conjunto inicial de dados brutos é reduzido a grupos mais gerenciáveis
  - Utilizado em processamento de imagens, textos em linguagem natural
- Feature Construction
  - Construção manual de novas features a partir de dados brutos
  - Derivação baseada nas features existentes
  - Requer muito tempo para estruturar e analisar
  - Ex: Combinação, Agregação, decomposição
  - Feature Learning
- Processo Iterativo
  - Brainstorm
    - Entender o problema, analisar os dados, estudar e propor o conjunto de features
  - Criar
    - Usar feature extraction automáticas e criação de features manuais (ou os dois)
  - Seleção
    - Usar feature importance e selection, crie visões para os modelos
  - Avaliar
    - Estimar a precisão (qualidade) dos modelos usando as features selecionadas

Principais Técnicas

- Extracting Date
  - Extrair parte das datas em diferentes colunas
  - Extrair diferença entre data atual e as colunas em termos de anos, meses, dias etc
  - Extrair alguns recursos específicos da data: Nome do dia da semana, Fim de semana ou não, feriado ou não, etc
- Imputation
  - Técnica para tratamento de valores nulos visto que esses valores afetam o desempenho do modelo
  - A solução mais simples seria descartar as linhas ou a coluna inteira.
  - Numerical Imputation
    - Aplicar tratamento para preencher os campos nulos
    - Exceto que se tenha um valor padrão para valores ausentes, o ideal é utilizar as medianas das colunas. Médias são sensível a pontos extremos.
  - Categorical Imputation
    - Para valores uniformes pode-se criar uma categoria Outros ou Não Informado
    - Uma solução também utilizada é substituir valores ausentes pela categoria que houver mais ocorrência
- Handling Outliers
  - Melhor maneira de identificar Outliers é através de visualização de dados
  - Identificação através de estatística é menos precisa, porém mais rápida
  - Os Outliers podem ser removidos ou limitados. Se limitados podem afetar a distribuição dos dados enquanto que se removidos afetam a quantidade de dados.
  - Outlier Detection with Standard Deviation
    - Tratar valores que estiverem muito distante do desvio padrão
    - Solução simples é utilizar valores entre 2 e 4
    - Pode ser gerado um Standard Score que padroniza a distância entre um valor e a média usando o desvio padrão
  - Outlier Detection with Percentiles
    - Assumir uma certa porcentagem do valor de cima ou de baixo como um erro externo
    - O percentual deve ser definido através de análise da distribuição dos dados
- One-Hot Encoding
  - Espalha valores de uma coluna em várias colunas e atribui os valores 0 e 1.
  - Altera dados categóricos para numéricos sem perder informação
- Grouping Operations
  - Transformar uma dataset muito grande em um menor
  - Categorical Column Grouping
    - Técnica 1: Selecionar os dados com maior frequência através de função lambda
    - Técnica 2: Espalha valores de uma coluna em várias colunas porém de forma agregada
    - Técnica 3: Utilizar a técnica 1 porém após a técnica One-Hot Encoding, preservando todos os dados
  - Numerical Column Grouping
    - Agrupamentos tradicionais de números: Soma, Média, Proporção
- Feature Split
  - Não existe regra fixa, depende basicamente das características da feature
  - Organizar dados do tipo string para facilitar a compreensão
  - Permite agrupar dados que não eram agrupáveis
  - Exemplo de dados que podem sofrer essa técnica:
    Leandro Carnevali Ferro -> Leandro / Ferro
    Toy Story (1995) -> Toy Story / 1995
- Scaling
  - Normalization
    - Dimensiona os valores entre 0 e 1.
    - Antes de utilizar essa técnica é interessante usar técnica Handling Outliner, visto que serão considerados os valores de MAX e MIN.
    - Norm = (X -X(min)) / (X(max) -X(min))
  - Standardization
    - Escala os valores levando em consideração o desvio padrão
    - Diminui efeito dos Outliers
    - Standardization (Z Score) = (X - Média) / Desvio Padrão
  - Utilizado em dados numéricos para pegar features de diferentes grandezas, exemplo kg e metros e colocar na mesma escala.
- Log Transform
  - Log(X+1)
  - É uma das mais usadas em engenharia de feature
  - Ajuda a lidar com dados assimétricos e, após a transformação, a distribuição se torna mais aproximada ao normal.
  - Diminui o efeito dos outliers, devido à normalização das diferenças de magnitude.
    Ex. 15 a 20 anos tem magnitude diferente de 60 a 65 anos
  - Os dados aplicados à transformação de log devem ter apenas valores positivos
- Binning
  - Ao selecionar algo, sacrifica informações e torna seus dados mais regularizados.
  - Pode ser aplicado em dados Categóricos e Numéricos
  - Objetivo é ter um modelo mais robusto e evitar overfitting, no entanto, isso tem um custo para o desempenho.
  - Ex: Países podem virar regiões
    Peso pode ser Magro, Normal, Obeso
  - A troca entre desempenho e overfitting é o ponto principal do processo de binning.
  - É interessante aplicar essa técnica em rótulos com baixa frequência, exemplo "Outros", para melhorar a robustez do modelo

Tipos de Aprendizado de Máquina

Online Learning

- Cobre todos os tipos de aprendizado de máquina
- Ajusta o modelo de forma iterativa conforme mais dados são carregados no modelo

Offline Learning

- Cobre todos os tipos de aprendizado de máquina
- Os dados já foram trabalhados e utilizados para o treino e os dados novos são somente para efetuar as previsões

Tipos de Algoritmos de Aprendizado de Máquina

Aprendizado por Reforço

- Não existe conjunto de treinamento, rotulado ou não
- Baseado em tentativa e erro
- Algoritmos
  - Contextual Bandits
  - K-armed bandits
- Agentes
  - Se relaciona com o Ambiente através de ações
- Estado
  - O estado do ambiente é atualizado a cada ação do agente
- Reforço (Recompensa)
  - Cada ação tem um nível de castigo ou recompensa
  - Maior a recompensa em relação a ação que deixa mais próximo o agente do objetivo
  - A recompensa é dada logo após o resultado da ação em alguns algoritmos, em outros, somente após a exploração de todo o ambiente.
- Ambiente
  - É o ambiente no qual está inserido o agente
  - O ambiente dá uma resposta a ação do agente
- Ação
  - São as ações que o agente consegue praticar
  - Algumas mudanças no ambiente podem permitir novas ações

Supervisionado

- Classes de Algoritmos
  - Classificação
    - Classificação compreende a uma tarefa de Aprendizado Não Supervisionado, similar aos algoritmos de Clustering
    - Utilizado quando o rótulo vem de um conjunto finito e não ordenado
  - Regressão
    - Quando o rótulo é um número real e não uma característica
- Semi Supervisionado
  - Utilizado quando não se consegue rotular todos os seus dados
  - Aprende utilizando dados rotulados e não rotulados
  - Este método permite-nos melhorar significativamente a acurácia por usar dados rotulados e não rotulados
- Aprendizado Ativo
  - O algoritmo aprende através de iteração humana para identificação dos rótulos
- Aprendem a executar uma dada tarefa a partir de exemplos rotulados
- Interativamente o algoritmo faz previsões e aprimora sua função até atingir um nível aceitável de acerto
- Exemplos de Algoritmos
  - Regressão Linear
  - Árvore de Decisão
  - Random Forest
  - Gradiente Boosting Machines
  - Support Vector Machines

Não Supervisionado

- Utilizado como pré-processamento para identificar características importantes para melhorar Regressão/Classificação
- Encontrar features relevantes sobre o dado de forma automática
- Permite detectar anomalias automaticamente, ou seja, valores que destoam do conjunto de dados
- Em visualização de dados, permite encontrar padrões visuais em coleções com muitas observações de features diferentes
- Utiliza dados não rotulados para aprender características a partir de padrões semelhantes
- Alguns objetos podem diferir largamente de todos os grupos e, deste modo, podemos detectar anomalias
- Classes de Algoritmos
  - Sumarização
  - Agrupamento
  - Associação

Aprofundamento de Conceitos

- Erro de Treino
  - Erro empírico: Erro do modelo nos dados de treino
- Erro de Teste
  - Erro de validação: Erro do modelo nos dados de teste
- Overfitting
  - Modelo excessivamente complexo para um dado conjunto de dados
  - Faz com que a entrada de treino seja memorizada ao invés de aprendida
- Underfitting
  - Ocorre quando um modelo é muito simples e não consegue generalizar todas as características dos dados
- Generalização
  - É a habilidade do modelo de predizer corretamente instâncias ainda não vistas
- Ruído
  - Dados indesejados presentes na coleção que podem gerar efeitos inesperados para o modelo
- Baseline
  - Base de comparação para a eficácia de um algoritmo
- Variância
  - Como o modelo reage a novas entradas de treino
  - Modelos com alta variância podem mudar suas predições em resposta à nova composição da entrada de treino (Overfitting)
- Viés
  - A habilidade do modelo de generalizar os dados, portanto, um viés alto é relacionado ao underfitting
- Bias Trade-off
  - É o ponto no modelo on se minimiza simultaneamente os efeitos de viés e variância que impedem que os algoritmos generalizem bem