

Tarefas e Ferramentas Avançadas em Machine Learning

Objetivos da aula

- A Tarefa de Classificação
- Encontrando Grupos com Agrupamento
- O Poder das Redes Neurais



O que é Classificação?

● Definição

- A tarefa de atribuir um rótulo (classe) a um objeto com base em suas características. É um aprendizado

● Como Avaliamos o Sucesso?

- Precisão: Importante para evitar falsos resultados.
- Recall: Importante para não deixar passar nada despercebido.
- F1-Score: Um equilíbrio inteligente entre Precisão e Recall.



Ferramentas Modernas: Ensembles

● A Ideia Central

A sabedoria coletiva. Em vez de confiar em um único "especialista" (um modelo), combinamos as previsões de vários.

● Como Funciona?

- Usando Árvores de Decisão, que são modelos que fazem uma série de perguntas "sim/não" para chegar a uma conclusão.

● Os Campeões Modernos

XGBoost e LightGBM:



Classificação vs. Agrupamento

- **Classificação (Supervisionado)**

- Você sabe as classes de antemão.
- Os dados de treino têm rótulos (respostas).
- Objetivo: Treinar um modelo para rotular novos dados.

- **Agrupamento (Não Supervisionado)**

- Você não sabe quais são os grupos.
- Os dados não têm rótulos.
- Objetivo: Encontrar grupos "naturais" que existem nos dados.

Como Encontrar Grupos?

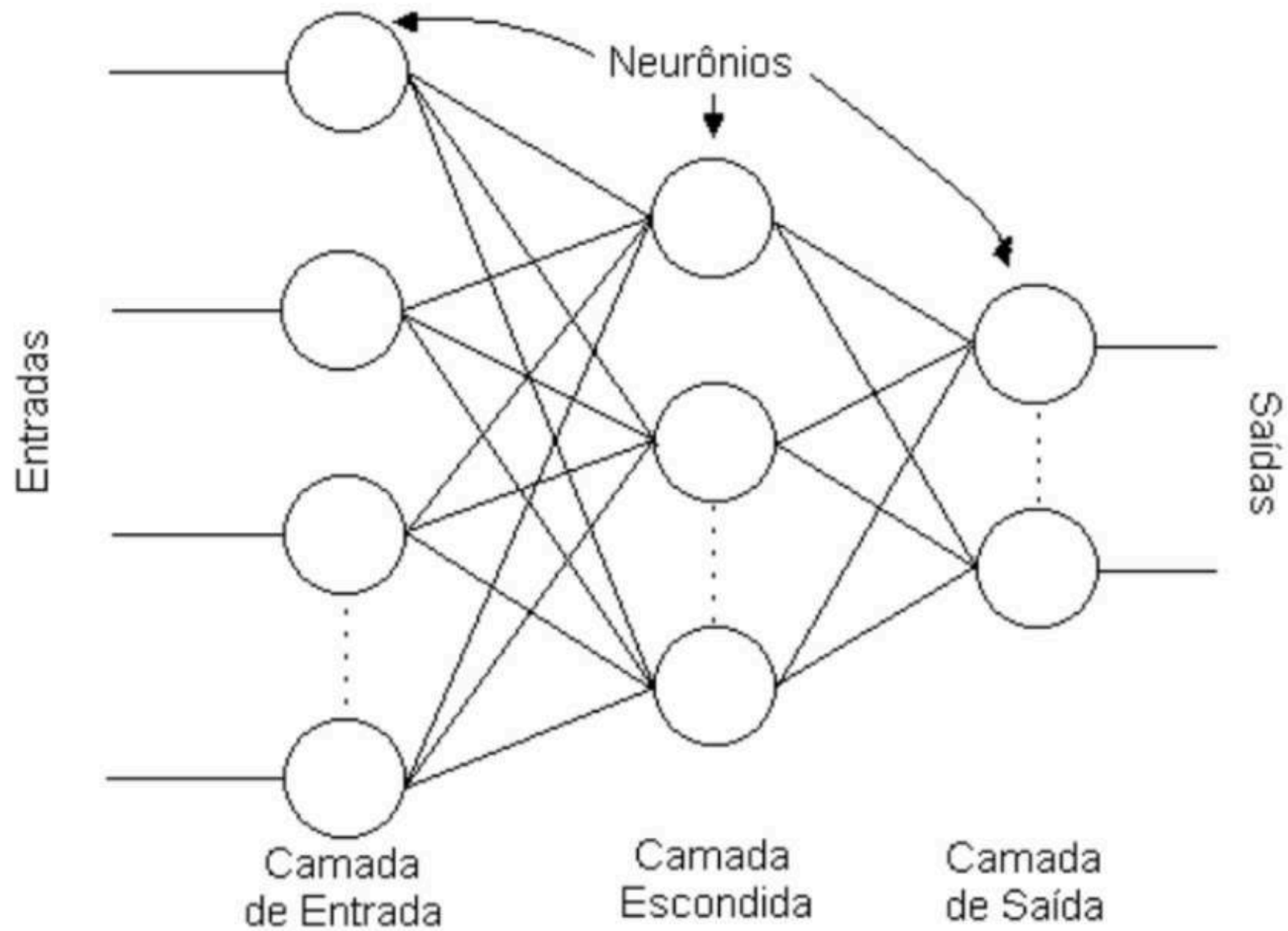
● K-Means

- Você define de antemão quantos grupos ('K') quer encontrar. O algoritmo então acha os centros desses grupos. É rápido, mas assume que os grupos são circulares.

● DBSCAN

- Não precisa definir o número de grupos. Ele encontra áreas de "alta densidade" e as define como grupos, sendo ótimo para formatos irregulares e para identificar outliers.





O Salto para o Deep Learning

● O que é "Deep"?

- Simplesmente usar Redes Neurais com muitas camadas ocultas. Isso permite aprender padrões incrivelmente complexos.

● CNNs (Redes Neurais Convolucionais)

- Os "especialistas em linguagem e sequências".
- Usam um "mecanismo de atenção" para focar nas partes mais importantes da entrada. São a base de modelos como o GPT.

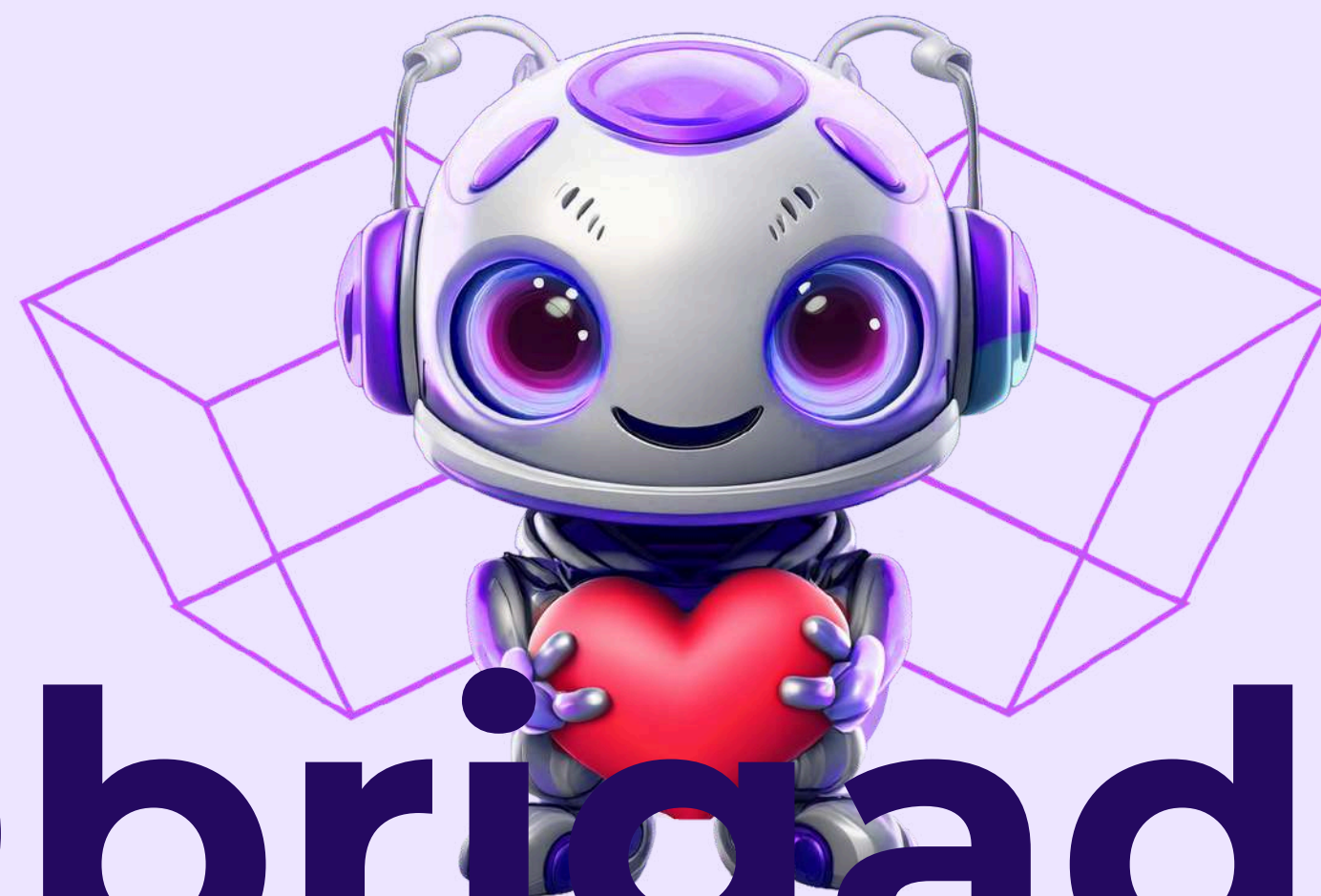
● Transformers

- Os "especialistas em linguagem e sequências".
- Usam um "mecanismo de atenção" para focar nas partes mais importantes da entrada. São a base de modelos como o GPT.

Recapitulação do que foi visto

- **Classificação:**
 - Tarefa supervisionada para atribuir rótulos. Medimos com Precisão e Recall.
- **Agrupamento:**
 - Tarefa não supervisionada para encontrar grupos. Avaliamos com o Índice de Silhueta.
- **Redes Neurais:**
 - Uma família de modelos poderosos que podem ser usados para ambas as tarefas (e mais), especialmente em sua forma Deep para dados complexos.





Obrigada!

Perguntas?