

Roteiro de apresentação

Tarefa PIBIC

Daniela Lima
2023

SUMÁRIO

1	APRENDIZADO SUPERVISIONADO E NÃO SUPERVISIONADO	2
1.1	DIFERENÇAS ENTRE APRENDIZADO SUPERVISIONADO E NÃO SUPERVISIONADO	2
1.1.1	<i>Aprendizagem supervisionada</i>	<i>2</i>
1.1.2	<i>Aprendizagem não supervisionada</i>	<i>2</i>
1.1.3	<i>Extra</i>	<i>3</i>
1.1.3.1	Algoritmos de cluster	3
1.1.3.2	Agrupamento hierárquico	3
1.1.3.3	Agrupamento de k-médias	4
1.1.3.4	Modelos de mistura gaussiana	4
1.1.3.5	Agrupamentos difusos	4
1.1.3.6	Associação	5
1.1.3.7	Modelos de variáveis latentes	5
1.1.4	<i>Utilização</i>	<i>5</i>
1.1.4.1	Aprendizagem não supervisionada	5
1.1.4.2	Aprendizagem supervisionada	6
2	DEEP LEARNING	8
2.1	MACHINE LEARNING VS DEEP LEARNING	8
2.1.1	<i>Machine Learning</i>	<i>8</i>
2.1.2	<i>Deep Learning</i>	<i>8</i>
2.1.3	<i>As principais diferenças entre Machine Learning e Deep Learning</i>	<i>9</i>
2.2	REDE NEURAL CONVOLUCIONAL	10
2.2.1	<i>Benefícios do uso de CNNs para aprendizado profundo</i>	<i>11</i>
2.3	CONVOLUÇÕES	11
2.4	FILTRO	12
2.4.1	<i>Como é decidido o tamanho do filtro?</i>	<i>13</i>
2.4.2	<i>Volumes</i>	<i>13</i>
2.5	COMO UMA IMAGEM É TRANSFORMADA EM UMA CNN	14
2.6	CUSTO COMPUTACIONAL DE UMA CNN	14
3	FONTES	16
3.1	EXTRAS	17

1 APRENDIZADO SUPERVISIONADO E NÃO SUPERVISIONADO

1.1 DIFERENÇAS ENTRE APRENDIZADO SUPERVISIONADO E NÃO SUPERVISIONADO

O aprendizado não supervisionado possui dados não rotulados que o algoritmo deve tentar entender por conta própria. O aprendizado supervisionado é onde os conjuntos de dados são rotulados para que haja uma resposta com a qual a máquina pode medir sua precisão.

1.1.1 APRENDIZAGEM SUPERVISIONADA

Se um computador estivesse aprendendo a identificar frutas em um ambiente de aprendizado supervisionado, ele receberia imagens de exemplo de frutas que seriam rotuladas. Isso é chamado de dados de entrada. Por exemplo, os rótulos diriam que as bananas são longas, curvas e amarelas, as maçãs são redondas e vermelhas, enquanto uma laranja é esférica, com aparência de cera e da cor laranja. Depois de algum tempo, a máquina deve ser capaz de identificar com segurança qual fruta é qual, com base nesses descritores.

1.1.2 APRENDIZAGEM NÃO SUPERVISIONADA

Em contraste, a aprendizagem não supervisionada ocorre quando não há categorização ou rotulação dos dados. A máquina não terá ideia do conceito de fruta, por isso não pode etiquetar os objetos. No entanto, pode agrupá-los de acordo com suas cores, tamanhos, formas e diferenças. A máquina agrupa as coisas de acordo com semelhanças, encontrando estruturas e padrões ocultos em dados não rotulados. Não existe caminho certo ou errado, e nenhum professor. Não há resultados, apenas uma análise pura dos dados.

É permitir que a máquina entenda os dados e os processe como bem entender.

Esse tipo de aprendizagem é usado para explorar dados desconhecidos. Ela pode revelar padrões que podem ter sido perdidos ou examinar grandes conjuntos de dados que seriam demais para serem resolvidos por um ser humano.

1.1.3 EXTRA

A aprendizagem não supervisionada usa uma variedade de algoritmos para ajustar os dados em grupos amplos, agrupamentos e associações.

1.1.3.1 ALGORITMOS DE CLUSTER

O Clustering ou agrupamento é quando os objetos são agrupados em subconjuntos chamados clusters. Esta é uma das melhores maneiras de obter uma visão geral da estrutura de seus dados. Haverá algumas características semelhantes nesses grupos. Este método é projetado para ter grupos com as mesmas características e, em seguida, atribuí-los aos clusters relevantes.

1.1.3.2 AGRUPAMENTO HIERÁRQUICO

É quando a máquina agrupa coisas que vão juntas em uma árvore de cluster. Todos os dados são um cluster, então eles se dividem em clusters cada vez menores. Os dados pertencerão a um conjunto em cascata de clusters, desde os mais genéricos até os mais específicos e bem agrupados. Portanto, o resultado é que você vê como os diferentes subgrupos se relacionam uns com os outros ou quão distantes eles estão.

1.1.3.3 AGRUPAMENTO DE K-MÉDIAS

Este algoritmo separa os dados em clusters distintos que não foram rotulados nos dados. A distância ao centro do cluster depende da força da associação. Os pontos de dados podem pertencer a apenas um cluster. Um k maior significa um grupo menor com mais granularidade da mesma maneira. Cada cluster é atribuído a um rótulo de ponto de dados.

1.1.3.4 MODELOS DE MISTURA GAUSSIANA

Com base em uma distribuição normal da curva de sino, os agrupamentos de grupos são espalhados em densidades normais esperadas, mostrando subpopulações nos dados gerais.

1.1.3.5 AGRUPAMENTOS DIFUSOS

Esses clusters podem se sobrepor, portanto, cada ponto de dados pode pertencer a quantos clusters forem relevantes, em oposição ao agrupamento rígido, em que os pontos de dados só podem pertencer a um cluster. Este é o diagrama de Venn do mundo da aprendizagem não supervisionada.

O armazenamento em cluster pressupõe relacionamentos entre grupos e, portanto, nem sempre é a melhor forma de segmentação de clientes - esse algoritmo não trata os pontos de dados como indivíduos. Você precisa aplicar mais métodos estatísticos para analisar os dados mais detalhadamente.

1.1.3.6 ASSOCIAÇÃO

No machine learning, o algoritmo cria regras que encontram associações entre pontos de dados. Ele encontra as relações entre as variáveis, identificando itens que tendem a ocorrer juntos. Por exemplo, a análise de cestas em supermercados pode ver quais itens as pessoas tendem a comprar ao mesmo tempo - sopa e pão, por exemplo. Ou, quando as pessoas compram uma casa nova, o que é provável que comprem também novo? Este algoritmo é excelente para identificar oportunidades de marketing

1.1.3.7 MODELOS DE VARIÁVEIS LATENTES

A modelagem de variáveis latentes mostra a relação entre as variáveis observáveis (ou variáveis manifestas) e aquelas que estão ocultas ou não observadas (variáveis latentes). Modelos de variáveis latentes são usados principalmente no pré-processamento / limpeza de dados, para reduzir os recursos em um conjunto de dados ou dividir o conjunto de dados em vários componentes.

1.1.4 UTILIZAÇÃO

1.1.4.1 APRENDIZAGEM NÃO SUPERVISIONADA

O melhor momento para usar a aprendizagem não supervisionada é quando não há dados pré-existent sobre os resultados preferidos. O aprendizado não supervisionado pode identificar recursos que podem ser úteis na categorização de conjuntos de dados desconhecidos.

- Redução de dimensionalidade

Como o nome sugere, o algoritmo trabalha para reduzir as dimensões dos dados. É usado para extração de recursos.

Extrair os recursos importantes do conjunto de dados é um aspecto essencial dos algoritmos de machine learning. Isso ajuda a reduzir o número de variáveis aleatórias no conjunto de dados, filtrando recursos irrelevantes.

Benefícios:

1. Redução da complexidade dos modelos
 2. Aumento da precisão dos modelos
 3. Economiza recursos computacionais
 4. Redução do tempo de treinamento dos modelos
- Clustering

O Clustering ou agrupamento é quando os objetos são agrupados em subconjuntos chamados clusters. Esta é uma das melhores maneiras de obter uma visão geral da estrutura de seus dados. Haverá algumas características semelhantes nesses grupos. Este método é projetado para ter grupos com as mesmas características e, em seguida, atribuí-los aos clusters relevantes.

1.1.4.2 APRENDIZAGEM SUPERVISIONADA

- Classificação

A classificação refere-se a pegar um valor de entrada e mapeá-lo para um valor discreto. Em problemas de classificação, nossa saída geralmente consiste em classes ou categorias. Isso pode ser algo como tentar prever quais objetos estão presentes em uma imagem (um gato/um cachorro) ou se vai chover hoje ou não.

- Regressão:

(estimar valores)

A regressão está relacionada a dados contínuos. Em Regressão, os valores de saída previstos são números reais. Ele lida com problemas como prever o preço de uma casa ou a tendência do preço das ações em um determinado momento etc.

2 DEEP LEARNING

2.1 MACHINE LEARNING VS DEEP LEARNING

O machine learning usa algoritmos para tomar decisões com base no que aprendeu com os dados. Já o deep learning usa algoritmos em camadas para criar uma rede neural artificial que toma decisões inteligentes por conta própria.

2.1.1 MACHINE LEARNING

O Machine Learning (ML) é qualquer abordagem que emprega algoritmos para filtrar dados e encontrar padrões, similar a uma máquina que executa uma função mecânica específica. O algoritmo executa a função definida pelo engenheiro ou programador e analisa os dados para fornecer uma resposta.

Alguns exemplos de machine learning:

- Um aplicativo que encontrar oportunidades financeiras favoráveis;
- Máquinas inteligentes na Indústria 4.0;
- Um programa em busca de malware.

O machine learning é automatizado, mas só até certo ponto. No Machine Learning, o programador ainda precisa fornecer uma orientação. Caso o algoritmo dê uma previsão ruim ou incorreta, é necessária uma intervenção humana.

2.1.2 DEEP LEARNING

Qualquer menção à aprendizagem profunda será seguida pelo termo "redes neurais" – o conceito que liga as capacidades de processamento do cérebro humano ao Deep Learning.

Isso não é totalmente incorreto, mas essa explicação tende a exagerar as capacidades do deep learning.

O fato é que o Deep Learning funciona de forma semelhante ao ML, mas usando uma quantidade de dados maior. Ao invés de um ou dois algoritmos trabalhando ao mesmo tempo, o deep learning depende de um modelo mais sofisticado de algoritmos agrupados, chamado de rede neural artificial ou ANN (Artificial Neural Network). Com isso, seus resultados parecem mais “inteligentes” ou sofisticados.

É essa rede neural artificial que é inspirada nos nossos cérebros, onde cada algoritmo representa um neurônio. O conjunto desses neurônios analisa continuamente os dados e atualiza suas previsões, assim como o nosso cérebro constantemente recebe informações e tira conclusões. Exemplos de deep learning incluem a identificação de rostos a partir de fotos ou vídeos e o reconhecimento de uma palavra falada.

Uma grande diferença é que o Deep Learning, ao contrário do ML, se corrigirá no caso de uma previsão ruim, tornando a intervenção humana menos necessária. Se uma lâmpada, por exemplo, tivesse recursos de deep learning, ela poderia responder não apenas ao "escuro", mas também a frases como "não consigo ver" ou "onde está o interruptor de luz?".

2.1.3 AS PRINCIPAIS DIFERENÇAS ENTRE MACHINE LEARNING E DEEP LEARNING

Esta é uma pergunta comum e se você leu até aqui, provavelmente já sabe que não deve ser feita dessa forma. Algoritmos de deep learning são algoritmos de machine learning. Portanto, pode ser melhor pensar sobre o que torna o deep learning especial no campo do machine learning. A resposta: a estrutura do algoritmo ANN, a menor necessidade de intervenção humana e os maiores requisitos de dados.

Em primeiro lugar, enquanto os algoritmos tradicionais de machine learning têm uma estrutura bastante simples, como uma regressão linear ou uma árvore de decisão, o deep learning é baseado em uma rede neural artificial. Essa ANN multicamada é, como um cérebro humano, complexa e entrelaçada.

Em segundo lugar, os algoritmos de aprendizagem profunda requerem muito menos intervenção humana. Lembra do exemplo da Tesla? Se o reconhecimento de imagem do sinal STOP fosse um algoritmo de machine learning mais tradicional, um engenheiro de software escolheria manualmente os recursos e um classificador para classificar as imagens, verificaria se a saída é a necessária e ajustaria o algoritmo se não fosse o caso. No entanto, como um algoritmo de deep learning, os recursos são extraídos automaticamente e o algoritmo aprende com seus próprios erros (veja a imagem abaixo).

(img) O algoritmo de deep learning não precisa de um engenheiro de software para identificar recursos, mas é capaz de engenharia automática de recursos por meio de sua rede neural.

Em terceiro lugar, o deep learning requer muito mais dados do que um algoritmo tradicional de machine learning para funcionar corretamente. Machine Learning funciona com mil pontos de dados, deep learning muitas vezes apenas com milhões. Devido à complexa estrutura multicamadas, um sistema de deep learning precisa de um grande conjunto de dados para eliminar flutuações e fazer interpretações de alta qualidade.

2.2 REDE NEURAL CONVOLUCIONAL

Uma Rede Neural Convolutiva (ConvNet / Convolutional Neural Network / CNN) é um 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. O pré-processamento exigido em uma ConvNet é muito menor em comparação com outros algoritmos de classificação. Enquanto nos métodos primitivos os filtros são feitos à mão, com treinamento suficiente, as ConvNets têm a capacidade de aprender esses filtros / características.

A arquitetura de uma ConvNet é análoga àquela do padrão de conectividade de neurônios no cérebro humano e foi inspirada na organização do Visual Cortex. Os neurônios individuais respondem a estímulos apenas em uma região restrita do campo visual conhecida

como Campo Receptivo. Uma coleção desses campos se sobrepõe para cobrir toda a área visual.

2.2.1 BENEFÍCIOS DO USO DE CNNs PARA APRENDIZADO PROFUNDO

Para aplicativos de reconhecimento de imagem, classificação de imagem e visão computacional (CV), as CNNs são particularmente úteis porque fornecem resultados altamente precisos, especialmente quando muitos dados estão envolvidos. A CNN também aprende as características do objeto em iterações sucessivas à medida que os dados do objeto se movem pelas várias camadas da CNN. Esse aprendizado direto (e profundo) elimina a necessidade de extração manual de recursos (engenharia de recursos).

As CNNs podem ser retreinadas para novas tarefas de reconhecimento e construídas em redes preexistentes. Essas vantagens abrem novas oportunidades para usar CNNs para aplicações do mundo real sem aumentar as complexidades ou custos computacionais.

Como visto anteriormente, as CNNs são mais eficientes computacionalmente do que as NNs regulares, pois usam o compartilhamento de parâmetros. Os modelos são fáceis de implantar e podem ser executados em qualquer dispositivo, até mesmo em smartphones.

2.3 CONVOLUÇÕES

Matematicamente, a convolução é a soma do produto elementar de 2 matrizes.

Vamos considerar uma imagem 'X' e um filtro 'Y'. Ambos, ou seja, X e Y, são matrizes (imagem X está sendo expressa no estado de pixels). Quando convoluímos a imagem 'X' usando o filtro 'Y', produzimos a saída em uma matriz, digamos 'Z'.

Finalmente, calculamos a soma de todos os elementos em 'Z' para obter um número escalar, ou seja, $3+4+0+6+0+0+0+45+2 = 60$

As convoluções funcionam como filtros que enxergam pequenos quadrados e vão “escorregando” por toda a imagem captando os traços mais marcantes. Explicando melhor, com uma imagem e um filtro que cobre uma área da imagem com movimento de X saltos (chamado de stride), o filtro passará pela imagem inteira, por cada um dos canais, formando no final um feature map ou activation map.

2.4 FILTRO

Esses filtros se comportam como uma pequena janela ou região que realiza operações ao longo da imagem. Uma analogia comum é a luz de uma lanterna passando ao longo da imagem, região à região.

Em cada passo, cada elemento do filtro é multiplicado pelo elemento de mesma posição na região em que o filtro está sendo aplicado naquele instante. No fim, soma-se os resultados dessas multiplicações para ter um único valor como saída. Esse valor será o pixel correspondente na imagem filtrada.

O filtro vai “caminhando” ao longo da imagem realizando essa mesma operação. O resultado é uma imagem filtrada.

Como dito anteriormente, o objetivo da aplicação desses filtros é extrair características importantes na imagem. Porém, quando uma CNN é inicializada, os valores dos filtros são inicializados aleatoriamente, ou seja, eles inicialmente extraem características aleatórias. O objetivo de aplicar isso em uma rede neural é justamente que o computador aprenda por si só os valores ideais dos filtros. Isso significa que a rede aprende sozinha quais características da imagem são importantes para realizar a tarefa que lhe foi designada.

Uma única camada convolucional pode ser constituída de diversos filtros. Ao adicionar mais camadas (tornar a rede mais profunda), a ideia é que a rede realize a combinação de características extraídas por camadas anteriores. Por exemplo, se a primeira camada aprendeu a extrair bordas em diferentes orientações, a camada seguinte combina essas bordas para aprender a detectar formas mais complexas, como curvas e cantos.

2.4.1 COMO É DECIDIDO O TAMANHO DO FILTRO?

Por convenção, o valor de 'f', ou seja, o tamanho do filtro, geralmente é ímpar na visão computacional. Isso pode ser devido a 2 motivos:

- Se o valor de 'f' for par, podemos precisar de preenchimento assimétrico. Digamos que o tamanho do filtro, ou seja, 'f' é 6. Então, usando a equação 1, obtemos um tamanho de preenchimento de 2,5, o que não faz sentido.
- A segunda razão para escolher um filtro de tamanho ímpar, como um filtro 3×3 ou 5×5 , é que obtemos uma posição central e, às vezes, é bom ter um diferenciador.

2.4.2 VOLUMES

Um conceito importante no entendimento de uma CNN é o trabalho com volumes.

Uma imagem colorida pode ser considerada um volume, em que sua altura e largura correspondem à altura e à largura da imagem e sua profundidade corresponde ao número de canais de cor dessa imagem. Por ser colorida, normalmente refere-se aos canais de cor RGB (Red, Green, Blue), ou seja, essa imagem teria 3 canais e consequentemente o volume tem profundidade 3.

Uma imagem preto e branco tem profundidade 1, já que seu único canal de cor tem valores na escala de cinza.

Ao aplicar os filtros, eles devem ter profundidade igual ao que recebem como entrada. Filtros aplicados em imagens coloridas devem ter profundidade 3. Porém, em uma CNN, soma-se os resultados de cada canal, resultando em um canal só, ou uma imagem filtrada com altura e largura.

Os resultados da aplicação de todos os filtros da camada são diversas imagens filtradas, que quando organizados em uma camada, se tornam um novo volume com profundidade igual ao número de filtros aplicados (um canal resultante de cada filtro). Agora esse volume será a

entrada da próxima camada, o que significa que os filtros da próxima camada devem ter essa mesma profundidade.

Por isso as camadas foram representadas como volumes.

2.5 COMO UMA IMAGEM É TRANSFORMADA EM UMA CNN

Redes Neurais Convolucionais especializadas para aplicações em reconhecimento de imagem e vídeo. A CNN é usada principalmente em tarefas de análise de imagem, como reconhecimento de imagem, detecção e segmentação de objetos.

Existem três tipos de camadas em redes neurais convolucionais:

- Camada Convolucional: Essa etapa de extração de características tem como principal funcionamento a aplicação de filtros na imagem de entrada. A operação realizada por esses filtros é uma convolução, que dá nome à rede.
- Camada de Pooling: Usada para reduzir a dimensionalidade do mapa de recursos. Esse processo tende a tornar a CNN invariante a transformações geométricas, fazendo com que a rede seja capaz de dizer se há um objeto naquela imagem independentemente de onde ele esteja posicionado. Além disso, reduz consideravelmente o custo computacional da rede.
- Fully-Connected Layers: As camadas densas (totalmente conectadas) são conectados por pesos. Por fim, a camada de saída da rede será a previsão ou classificação.

2.6 CUSTO COMPUTACIONAL DE UMA CNN

“Mostramos que o aprendizado profundo não é computacionalmente caro por acidente, mas por design. A mesma flexibilidade que o torna excelente para modelar diversos fenômenos e superar modelos especializados também o torna drasticamente mais caro computacionalmente”

Para eles, um relatório do Synced estimou que o modelo de detecção de notícias falsas de Grover, da Universidade de Washington, custou US\$ 25.000 para ser treinado em cerca de duas semanas. A OpenAI supostamente acumulou US\$ 12 milhões para treinar seu modelo de linguagem GPT-3, e o Google gastou cerca de US\$ 6.912 treinando o BERT, um modelo de transformador bidirecional que redefiniu o estado da arte para 11 tarefas de processamento de linguagem natural.

Em um relatório de 2019, pesquisadores da Universidade de Massachusetts em Amherst concluíram que a quantidade de energia necessária para treinar e pesquisar um determinado modelo envolve a emissão de aproximadamente 626.000 libras de dióxido de carbono. Isso é equivalente a quase cinco vezes as emissões de um carro médio dos EUA durante toda a vida útil.

3 FONTES

[O que é aprendizagem não supervisionada?](#)

[IBM - Supervised vs. Unsupervised Learning: What's the Difference?](#)

[Supervised and Unsupervised Learning \[Differences & Examples\]](#)

[11 Different Uses of Dimensionality Reduction](#)

[DEEPLIZARD - Convolutional Neural Networks \(CNNs\) explained](#)

[3Blue1Brown - But what is a neural network?](#)

[3Blue1Brown - But what is a convolution?](#)

[DEEPLIZARD - Deep Learning Dictionary - Lightweight Crash Course](#)

[Deep Learning vs Machine Learning: entenda a diferença](#)

[Deep learning vs. machine learning: você sabe as diferenças?](#)

[Deep Learning vs. Machine Learning – What's The Difference?](#)

[IA Expert Academy - TensorFlow Playground \(Explicação sobre os neurônios\)](#)

[Entendendo Redes Convolucionais \(CNNs\)](#)

[Conhecendo a visão do computador: Redes Neurais Convolucionais](#)

[Image formation and Filtering](#)

[Convolutional Neural Network: An Overview](#)

[Convolutional neural network \(CNN\)](#)

[Universo Discreto - Aprenda FILTROS \(O que é Convolução e Filtros de Suavização\)](#)

[MIT researchers warn that deep learning is approaching computational limits](#)

3.1 EXTRAS

[Filters In Convolutional Neural Networks](#)

[3Blue1Brown - Gradient descent, how neural networks learn](#)

[Michael Nielsen's Neural Networks and Deep Learning Book](#)