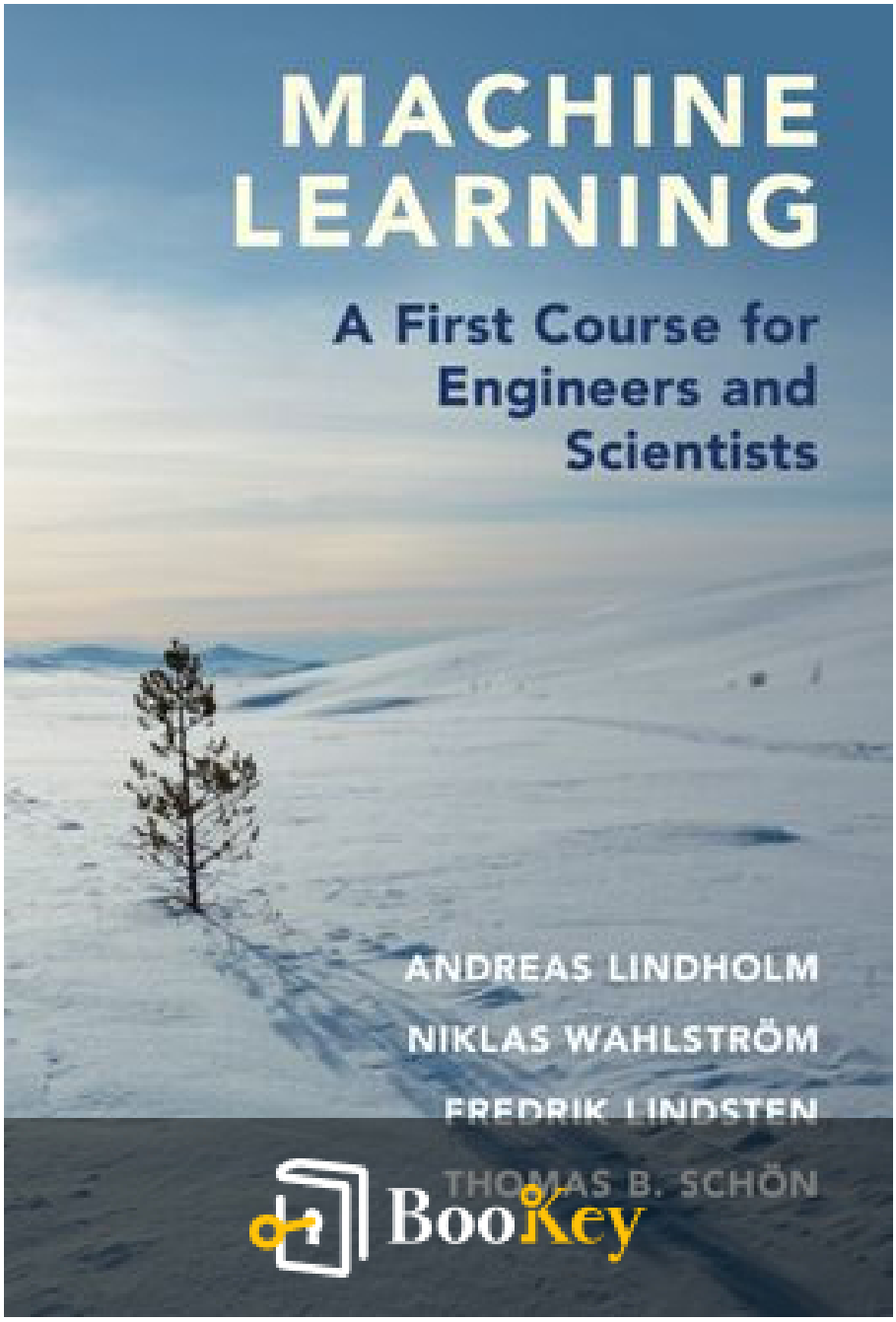
Aprendizado De Máquina PDF (Cópia limitada)

Andreas Lindholm



Teste gratuito com Bookey

Digitalize para baixar

Aprendizado De Máquina Resumo

Das Análises de Dados às Soluções Preditivas

Escrito por Books1

Teste gratuito com Bookey

Digitalize para baixar

Sobre o livro

Em uma era em que as máquinas não são mais meros observadores silenciosos, mas personagens ativas moldando as narrativas do nosso mundo, "Machine Learning" de Andreas Lindholm se apresenta como um convite fascinante para o transformador universo dos computadores que aprendem. Lindholm desmistifica os intrincados algoritmos que impulsionam nossa era digital, oscilando com elegância entre teorias fundamentais e práticas de ponta. Com uma clareza revigorante, este livro encapsula tanto um charme simples quanto uma sofisticação complexa, tornando o esotérico acessível tanto para novatos quanto para entusiastas experientes. Ele convida os leitores a olharem além do véu da programação convencional, prometendo uma jornada por paisagens onde os dados se tornam tanto a musa quanto o meio de inovações sem precedentes. Seja você um novato prestes a explorar sua primeira rede neural ou alguém que busca se aprofundar nas profundezas da inteligência de máquina, "Machine Learning" funciona como um mapa e uma bússola—guiando, informando e inspirando a cada passo.

Teste gratuito com Bookey

Digitalize para baixar

Sobre o autor

Andreas Lindholm é um especialista experiente na área de aprendizado de máquina, reconhecido por suas contribuições inovadoras e profunda compreensão da inteligência artificial. Com uma carreira que abrange mais de uma década, Lindholm tem sido fundamental na fusão de conceitos teóricos com aplicações práticas, dando passos significativos para aprimorar as capacidades da tecnologia. Ele possui diplomas avançados em ciência da computação e matemática aplicada, que serviram de base para sua paixão por aprendizado de máquina e soluções baseadas em dados. Como autor prolífico e palestrante requisitado em conferências internacionais, Lindholm influenciou tanto o meio acadêmico quanto os profissionais da indústria, sempre desafiando os limites do que é possível com aprendizado de máquina. Seu estilo de escrita acessível e perspicaz conquistou os leitores, ajudando a desmistificar teorias complexas para entusiastas e profissionais. Através de seu trabalho, Andreas Lindholm continua a inspirar a próxima onda de inovação no mundo da tecnologia.

Teste gratuito com Bookey

Digitalize para baixar

Ad

Experimente o aplicativo Bookey para ler mais de 1000 resumos dos melhores livros do mundo

Desbloqueie 1000+ títulos, 80+ tópicos

Novos títulos adicionados toda semana

duct & Brand | Liderança & Colaboração | Gerenciamento de Tempo | Relacionamento & Comunica

Estratégia de Negócios | Criatividade | Memórias | Conheça a Si Mesmo | Psicologia Po

mpreendedorismo | História Mundial | Comunicação entre Pais e Filhos | Autocuidado | Mi

Visões dos melhores livros do mundo

Teste gratuito com Bookey

Lista de Conteúdo do Resumo

Capítulo 1: Aprendizado supervisionado: uma primeira abordagem

Sure! The translation of "Chapter 2" into Portuguese is "Capítulo 2." If you need help with more text or specifics, feel free to provide the content!: Modelos paramétricos básicos e uma perspectiva estatística sobre a aprendizagem.

Capítulo 3: Claro! A tradução para o português da expressão "Understanding, evaluating and improving the performance" poderia ser:

"Compreender, avaliar e melhorar o desempenho."

Se precisar de mais ajuda, estou à disposição!

Sure! The translation of "Chapter 4" into Portuguese is "Capítulo 4." If you have more content or additional sentences to translate, feel free to share!: Modelos paramétricos em aprendizagem

Capítulo 5: Redes neurais e aprendizado profundo

Capítulo 6: Métodos de conjunto: Bagging e boosting

Capítulo 7: Transformações não lineares de entrada e núcleos

Capítulo 8: A abordagem bayesiana e os processos gaussianos

Certainly! Here's the translation of "Chapter 9" into Portuguese:

Teste gratuito com Bookey

Digitalize para baixar

Capítulo 9

If you need anything else or more text to be translated, feel free to ask!: Modelos generativos e aprendizado a partir de dados não rotulados.

Capítulo 10: Aspectos do usuário em aprendizado de máquina

Capítulo 11: Ética na aprendizagem de máquina

Teste gratuito com Bookey

Digitalize para baixar

Capítulo 1 Resumo: Aprendizado supervisionado: uma primeira abordagem

Capítulo 2: Aprendizado Supervisionado – Uma Primeira Abordagem

Neste capítulo, o conceito de aprendizado supervisionado em aprendizado de máquina é introduzido juntamente com dois métodos fundamentais: k-Vizinhos Mais Próximos (k-NN) e árvores de decisão. Esses métodos são intuitivos e diretos, mas servem como uma base sólida para entender técnicas mais sofisticadas.

2.1 Aprendizado de Máquina Supervisionado

O aprendizado supervisionado lida com dados que têm variáveis de entrada (frequentemente chamadas de características) e variáveis de saída correspondentes (rótulos). O objetivo é aprender a mapear as entradas para as saídas usando um modelo matemático, que pode então prever a saída para novos dados não vistos. Os dados de treinamento consistem em vários exemplos independentes, onde cada entrada é combinada com uma saída conhecida.

A distinção entre variáveis numéricas e categóricas é vital. Variáveis numéricas têm uma ordem inerente (como velocidade ou temperatura),

Teste gratuito com Bookey

Digitalize para baixar

enquanto variáveis categóricas não têm (como cores ou nomes). Essa diferença resulta em dois tipos de problemas: regressão (saída numérica) e classificação (saída categórica).

Exemplos Práticos:

1. **Classificando Músicas**: Imagine um aplicativo que identifica se uma música é de The Beatles, Kiss, ou Bob Dylan com base em características de áudio. Aqui, a saída é categórica, caracterizando um problema de classificação.

2. **Distâncias de Parada de Carros**: Dado dados sobre velocidades de carros e distâncias de parada, o desafio é prever as distâncias de parada para velocidades que não foram medidas. Como a saída (distância) é numérica, isso configura um problema de regressão.

Os modelos de aprendizado de máquina visam generalizar além dos dados de treinamento, o que significa que devem fazer previsões confiáveis para novas entradas. Um modelo que se ajusta perfeitamente aos dados de treinamento, mas falha em dados não vistos, é dito estar superajustado. A generalização é essencial para a aplicação prática de um modelo.

2.2 Um Método Baseado em Distância: k-NN

Teste gratuito com Bookey

Digitalize para baixar

O método k-NN prevê a saída com base nos exemplos mais próximos dos dados de treinamento, utilizando a distância euclidiana entre as variáveis de entrada. Para classificação, emprega uma votação da maioria entre os vizinhos mais próximos, enquanto que para regressão, utiliza a média deles.

Escolher o número correto de vizinhos (k) envolve uma troca entre flexibilidade e precisão. Muito poucos vizinhos podem levar ao superajuste (capturando ruídos em vez do padrão subjacente), enquanto muitos podem simplificar demais, perdendo padrões complexos.

Uma etapa importante é normalizar os dados de entrada para garantir que cada variável contribua igualmente para o cálculo da distância. A normalização pode ser alcançada redimensionando as entradas para uma faixa comum ou padronizando-as para que tenham uma média de zero e um desvio padrão de um.

2.3 Um Método Baseado em Regras: Árvores de Decisão

As árvores de decisão classificam os dados fazendo uma série de perguntas de sim/não, levando a um conjunto de regras que segmentam os dados em regiões distintas com previsões uniformes. Estruturalmente, são árvores binárias com condições (divisões) nos nós e previsões (nós folha) formando uma superfície de saída constante em partes.

Teste gratuito com Bookey

Digitalize para baixar

Para regressão, as divisões minimizam o erro de previsão, enquanto que para classificação, buscam aumentar a pureza dentro dos nós utilizando medidas como taxa de classificação incorreta, índice de Gini ou entropia. Essas medidas avaliam o quão bem cada divisão separa as classes.

Decidir a profundidade de uma árvore é crucial: mais profundidade aumenta a complexidade e o risco de superajuste, enquanto menos profundidade pode perder padrões. Técnicas como pré-definir uma profundidade máxima ou poda (remover partes da árvore) ajudam a controlar o tamanho da árvore.

Resumo e Leitura Adicional

O capítulo apresenta k-NN e árvores de decisão como métodos básicos, mas ilustrativos, para tarefas de aprendizado supervisionado. Embora intuitivos, eles servem como um passo inicial para entender modelos mais profundos e complexos. Leituras adicionais incluem trabalhos de Cover e Hart sobre k-NN e a obra de Breiman et al. sobre árvores de decisão, que fornecem uma compreensão mais abrangente desses métodos fundamentais.

Teste gratuito com Bookey

Digitalize para baixar

Sure! The translation of "Chapter 2" into Portuguese is "Capítulo 2." If you need help with more text or specifics, feel free to provide the content! Resumo: Modelos paramétricos básicos e uma perspectiva estatística sobre a aprendizagem.

Resumo do Capítulo 3: Modelos Paramétricos Básicos e uma Perspectiva Estatística sobre Aprendizado

Introdução à Modelagem Paramétrica

Este capítulo explora a modelagem paramétrica, um método utilizado em aprendizado de máquina supervisionado para resolver problemas aprendendo parâmetros do modelo a partir de dados de treinamento. O foco está na regressão linear e na regressão logística, ambos exemplos clássicos de modelos paramétricos, onde um conjunto de parâmetr aprendido e aplicado diretamente a novos dados sem necessidade de dados de treinamento uma vez que o modelo está finalizado.

3.1 Regressão Linear

A regressão, uma das tarefas centrais em aprendizado supervisionado, envolve prever uma saída numérica Ø5Üf a partir de v regressão linear assume que essa saída pode ser modelada como uma combinação afim (essencialmente uma combinação linear mais uma constante) de variáveis de entrada mais um termo de

Teste gratuito com Bookey

Digitalize para baixar

contabiliza o erro aleatório. O modelo é representado por:

\[Ø5Üf = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ld
\epsilon \]

Esta seção detalha como os parâmetros Ø5ß , incluind
determinados usando dados de treinamento por meio de métodos como a abordagem de mínimos quadrados. As "equações normais" formam a base para determinar os valores dos parâmetros, e quando a matriz \(X^T X \) é invertível, uma solução em forma fechada para Ø5ß e
Ilustrações como o modelo de distância de parada de um carro demonstram como a regressão linear é aplicada.

Funções de Perda e Funções de Custo
A otimização da regressão linear normalmente envolve minimizar uma função de custo derivada de uma função de perda, sendo a perda de erro quadrático a mais comum. Isso captura a diferença entre valores previstos e reais, motivando a abordagem da função de custo de mínimos quadrados.

3.2 Classificação e Regressão Logística
A classificação se distingue da regressão ao prever resultados categóricos. A regressão logística modifica o modelo de regressão linear para atender à classificação ao introduzir a função logística, que transforma saídas lineares em probabilidades. Para problemas de classificação binária, a regressão

Teste gratuito com Bookey
Digitalize para baixar

logística modela a probabilidade de uma classe usando:

\[g(x) = \frac{e^{\theta^T x}}{1 + e^{\theta^T x}} \]

Os parâmetros da regressão logística são aprendidos através da estimativa de máxima verossimilhança em vez de usar equações normais. Esse processo envolve otimizar uma função de custo baseando-se na perda de entropia cruzada, especialmente adequada para tarefas de classificação.

Predições e Limites de Decisão
Modelos logísticos prevêem categorias selecionando a classe mais provável, geralmente usando um limiar (por exemplo, 0,5 para resultados binários) para determinar os limites de decisão, que são lineares neste caso.

Extensões para Situações Multiclasse
Para classificação multiclasse, o modelo logístico é generalizado usando a função softmax, permitindo que o modelo produza uma distribuição de probabilidade sobre múltiplas classes. A função de custo empregada é a perda de entropia cruzada multiclasse.

3.3 Regressão Polinomial e Regularização
Para melhorar a flexibilidade do modelo, transformações não lineares, como a regressão polinomial, podem ser aplicadas, onde a por suas potências superiores (por exemplo, \(x^2, x^3 \)). No entanto, isso

Teste gratuito com Bookey
Digitalize para baixar

aumenta o risco de overfitting, onde os modelos capturam ruídos em vez de padrões. Técnicas de regularização, como a regulariz
Ridge), são introduzidas para combater o overfitting penalizando grandes magnitudes de parâmetro na função de custo, expressa como:

\[\text{Função de custo com regularização} = \frac{1}{n} ||X\theta - y||^2_2 + \lambda ||\theta||^2_2 \]

3.4 Modelos Lineares Generalizados
Generalizando ainda mais o conceito, modelos lineares generalizados (GLMs) adaptam modelos lineares a vários tipos de dados de saída, incorporando diferentes funções de ligação e distribuições além dos reinos binário e contínuo—como dados de contagem usando a regressão de Poisson, onde a ligação assegura a positividade do parâmetro médio, modelado como \(\lambda = \exp(\theta^T x) \).

Conclusão
Modelos paramétricos oferecem uma abordagem estruturada para aprender relações a partir dos dados, possibilitando tanto a previsão quanto a interpretação. Eles conectam relações lineares simples e distribuições de dados complexas, estabelecendo uma base para modelagens mais complexas abordadas em capítulos subsequentes.

Teste gratuito com Bookey
Digitalize para baixar

Capítulo 3 Resumo: Claro! A tradução para o português da expressão "Understanding, evaluating and improving the performance" poderia ser:

"Compreender, avaliar e melhorar o desempenho."

Se precisar de mais ajuda, estou à disposição!

Este capítulo aprofunda-se na compreensão, avaliação e aprimoramento do desempenho de modelos de aprendizado de máquina supervisionados. Encontramos diferentes metodologias para aprendizado de máquina, com o conceito central de que os modelos são adaptados aos dados de treinamento na esperança de que gerem previsões precisas para novos dados desconhecidos. O foco aqui é explorar quão viável é esperar que os modelos mantenham seu desempenho quando aplicados a conjuntos de dados não familiares, oferecendo insights sobre ferramentas práticas para avaliação e melhoria de modelos.

Erro Esperado em Novos Dados: Desempenho em Produção

O capítulo introduz o conceito de uma função de erro essencial para avaliar tarefas de classificação ou regressão. Essa função mede quão bem as previsões (Ø5ÜOØ5Üf(Ø5Üe)) se alinha (Ø5Üf). Métricas de erro comuns incluem a taxa de er

Teste gratuito com Bookey

Digitalize para baixar

tarefas de classificação e o erro quadrático para tarefas de regressão. É importante notar que a função de erro difere da função de perda utilizada no treinamento do modelo; enquanto a função de perda otimiza o aprendizado do modelo, a função de erro avalia o desempenho pós-treinamento.

O aprendizado supervisionado visa ajustar modelos para gerenciar efetivamente fluxos contínuos de dados, necessitando de um foco no desempenho em novos dados, descrito matematicamente pela função de erro médio. Isso envolve considerar entradas (x) e saídas (y) de maneira probabilística, uma tarefa desafiadora dada a complexidade e a natureza impraticável de definir a distribuição de dados do m

As previsões do modelo com base nos dados de treinamento (T) são expressas como Ø5ÜOØ5Üf(Ø5Üe; T). O erro esperado em derivado pela média sobre os possíveis pontos de dados de teste desconhecidos, enfatizando a capacidade do modelo de generalizar a partir de T. Além disso, o conceito de erro de treinamento, desempenho do modelo exclusivamente no conjunto de dados de treinamento, uma medida que geralmente não indica o desempenho em dados futuros.

Estratégias para Estimar Ø5Ü8new

Estimar com precisão Ø5Ü8new é crucial para avaliar

Teste gratuito com Bookey

Digitalize para baixar

da impossibilidade de calcular Ø5Ü8new diretamente d
dados desconhecidas, estimativas práticas como o erro de validação hold-out (Ø5Ü8hold-out) oferecem um substituto viável. A vali
dividir os dados disponíveis em conjuntos de treinamento e validação, uma abordagem que equilibra as necessidades de dados para treinamento e avaliação.

O método de validação cruzada em k-fold aprimora ainda mais a estimativa de Ø5Ü8new ao dividir os dados em k lotes, treinando
modelos em todos os lotes, exceto um, e validando no lote excluído. Esse processo iterativo aproveita todo o conjunto de dados, produzindo uma estimativa (Ø5Ü8k-fold) de Ø5Ü8new, além de permitir
hiperparâmetros.

Decomposição do Gap de Erro de Treinamento e Generalização

Decompor Ø5Ü8new em erro de treinamento e um gap d
a compreensão do desempenho do modelo. Normalmente, o desempenho de um modelo no treinamento supera sua capacidade de generalizar para novos dados, criando um gap de generalização que varia com a complexidade do modelo—modelos mais complexos tendem a se adaptar demais aos detalhes específicos dos dados de treinamento, correndo o risco de overfitting. Por outro lado, o underfitting ocorre quando os modelos não possuem complexidade suficiente para capturar padrões nos dados de maneira eficaz.

Teste gratuito com Bookey
Digitalize para baixar

Esse equilíbrio requer uma adaptação adequada, otimizando a complexidade do modelo para um Ø5Ü8new mínimo.

Decomposição de Viés e Variância

A decomposição de viés e variância oferece uma nova perspectiva na análise de Ø5Ü8new, segmentando-o em viés, variância e ruíd mede erros sistemáticos nas previsões, enquanto a variância indica a sensibilidade das previsões às variações nos dados. Aumentar a flexibilidade do modelo reduz o viés, mas pode amplificar a variância, destacando a importância de encontrar um equilíbrio ideal entre viés e variância. Esse equilíbrio é fundamental para minimizar Ø5Ü8new e a generalização robusta do modelo.

Ferramentas para Avaliar Classificadores Binários

Para avaliar classificadores binários de forma abrangente, o capítulo apresenta ferramentas como a matriz de confusão e a curva ROC, auxiliando na avaliação de desempenho além das taxas básicas de erro de classificação. Para problemas desequilibrados ou assimétricos—comuns em áreas como a detecção de doenças—métricas como precisão, recall e F1 são vitais. O PR-AUC complementa o ROC-AUC na avaliação das compensações entre precisão e recall do classificador de forma abrangente.

Teste gratuito com Bookey

Digitalize para baixar

Conclusões

Este capítulo ressalta a importância de metodologias de avaliação robustas no aprendizado supervisionado, fundamentais para desenvolver modelos que performem de maneira confiável em dados desconhecidos. Por meio da exploração detalhada das técnicas de estimativa de erro, análise de viés-variância e ferramentas de classificação binária, equipa os profissionais com uma compreensão holística das dinâmicas de desempenho dos modelos, estabelecendo uma base para tópicos mais avançados nos capítulos seguintes.

Seção	Resumo
Introdução do Capítulo	Este capítulo se concentra em entender e aprimorar o desempenho de modelos de aprendizado de máquina supervisionado, explorando ferramentas para avaliação e melhoria.
Erro Esperado em Novos Dados: Desempenho em Produção	Apresenta a função de erro para avaliar previsões. Discute a importância do desempenho em novos dados e faz a distinção entre erros de treinamento e erros em novos dados.
Estratégias para Estimar Ø5Ü8new	Descreve métodos como validação hold-out e validação cruzada k-fold para estimar o erro em novos d crucial para o aprimoramento do modelo.
Decomposição do Erro de Treinamento e da Lacuna de Generalização	Explica a lacuna de generalização entre o desempenho em dados de treinamento e novos dados, afetada pela complexidade do modelo e pelo risco de overfitting ou underfitting.
Decomposição de	Descreve a decomposição dos erros de previsão em viés,

More Free Book

undefined

Seção	Resumo
Viés e Variância	variância e ruído. Enfatiza a necessidade de um equilíbrio no trade-off entre viés e variância para minimiza
Ferramentas para Avaliação de Classificadores Binários	Destaque para ferramentas como matriz de confusão e curva ROC para avaliar classificadores binários, com métricas especiais para situações de dados desbalanceados.
Conclusões	Enfatiza o papel de metodologias de avaliação robustas para assegurar o desempenho confiável do modelo em dados não vistos, preparando o terreno para tópicos mais avançados.

More Free Book

undefined

Pensamento Crítico

Ponto Chave: Erro de Dados Novos Esperados: Desempenho em Produção

Interpretação Crítica: Imagine enfrentar cada novo dia como um modelo diante de dados desconhecidos. Este capítulo convida você a ver além das conquistas iniciais e a ouvir as lições do 'erro de dados novos esperados'. À medida que avança pela vida, não se trata apenas de se destacar em tarefas familiares, mas de se preparar para o inesperado com resiliência. Cultive uma mentalidade curiosa que abrace a incerteza e priorize experiências que ampliem seus horizontes. Lembre-se, assim como os modelos que descobrem dados não visíveis, você também pode prosperar adaptando-se, avaliando e refinando continuamente sua compreensão. Deixe este insight chave inspirá-lo a buscar crescimento através da exploração, garantindo que você permaneça versátil e perspicaz, assim como um modelo que visa generalizar de forma eficaz além de suas condições de treinamento.

Teste gratuito com Bookey

Digitalize para baixar

Sure! The translation of "Chapter 4" into Portuguese is "Capítulo 4." If you have more content or additional sentences to translate, feel free to share!: Modelos paramétricos em aprendizagem

Capítulo 5: Aprendendo Modelos Paramétricos

Este capítulo aprofunda-se no conceito de modelagem paramétrica, um componente fundamental de muitos métodos de aprendizado de máquina supervisionados. Ele elabora três tópicos centrais: funções de perda, regularização e otimização, elementos fundamentais ao lidar com modelos paramétricos como a regressão linear e a regressão logística. Embora discutidos brevemente no Capítulo 3, esses conceitos são explorados em detalhes neste capítulo, enfatizando sua aplicabilidade para além dos modelos paramétricos.

5.1 Princípios da Modelagem Paramétrica

A modelagem paramétrica, introduzida por meio da regressão linear e da regressão logística, pode ser estendida para abranger modelos mais complexos. O capítulo apresenta uma estrutura geral para modelos paramétricos, que se baseia no aprendizado de funções a partir de dados usando parâmetros adaptáveis, denotados por ¸. O ca

Teste gratuito com Bookey

Digitalize para baixar

funções não precisam ser lineares, permitindo relacionamentos mais elaborados e não-lineares, como o modelo de cinética de Michaelis-Menten para reações enzimáticas, caracterizado por sua dependência paramétrica não-linear.

Em seguida, o capítulo explora como esses modelos, que frequentemente assumem um ruído gaussiano aditivo, podem ser adaptados por meio de diferentes funções de perda, fornecendo meios flexíveis para descrever a relação entre entrada e saída. Por exemplo, a transição de modelos lineares envolve o uso de funções não-lineares arbitrárias para retratar relações de maneira mais precisa. Modelos não-lineares, como a regressão logística para classificação binária, podem ser expandidos para modelos multiclasses e além, utilizando funções softmax e redes neurais, que serão discutidas em capítulos posteriores.

Minimização da Perda e Generalização

Aprender um modelo paramétrico é essencialmente otimizar os parâmetros do modelo de forma que o erro de previsão ou a função de perda seja minimizada sobre os dados de treinamento. No entanto, o verdadeiro objetivo é a generalização—alcançar um modelo que preveja com precisão dados não vistos, o que o objetivo de treinamento apenas serve como um proxy. Para alinhar esses objetivos, várias estratégias são adotadas, incluindo considerar a precisão estatística em relação aos esforços computacionais,

Teste gratuito com Bookey

Digitalize para baixar

adotar uma função de perda distinta da função de erro para uma generalização mais robusta e empregar técnicas como parada antecipada e regularização explícita.

5.2 Funções de Perda e Modelos Baseados em Verossimilhança

Escolher uma função de perda apropriada é crítico, já que isso impacta o processo de aprendizado do modelo e sua eventual capacidade de generalização. Diferentes funções de perda destacam várias características nos modelos. Para regressão, as escolhas comuns incluem a perda de erro quadrático e a perda de erro absoluto, cada uma com diferentes implicações sobre a robustez do modelo, particularmente na presença de outliers. O conceito é estendido à classificação, onde funções como entropia cruzada, perda de hinge e perda logística definem várias abordagens para modelar probabilidades de classe, cada uma oferecendo vantagens e limitações únicas.

A adoção de modelos baseados em verossimilhança permite integrar suposições probabilísticas sobre os dados nos processos de aprendizado. Essa abordagem não apenas ajuda na seleção das funções de perda, mas também revela as conexões matemáticas entre distribuições de dados estruturados e parâmetros do modelo, levando a estruturas de modelo mais informadas.

Teste gratuito com Bookey

Digitalize para baixar

5.3 Regularização

A regularização é apresentada como uma técnica para evitar o overfitting, especialmente com modelos complexos. A regularização Œ2 (regressão ridge) e Œ1 (LASSO) são destacadas, introduzindo termos de penalização para manter pequenos valores de parâmetro, a menos que os dados indiquem fortemente o contrário. A regularização explícita frequentemente envolve a modificação da função de custo, melhorando a simplicidade do modelo, que tipicamente se correlaciona com uma melhor generalização. O capítulo também cobre métodos de regularização implícita, como a parada antecipada, que serve como um eficaz dissuasor contra o overfitting em contextos de otimização iterativa.

5.4 Otimização de Parâmetros

Um aspecto crítico do aprendizado de modelos paramétricos é a solução eficiente de problemas de otimização. Vários algoritmos, como gradiente descendente, método de Newton e descida de coordinate, são explorados quanto à sua aplicabilidade em diferentes configurações de problemas. Técnicas para gerenciar a taxa de aprendizado e a inicialização são essenciais, especialmente quando se trata de problemas não convexos, que são abordados com métodos como o método de Newton de região de confiança para resultados robustos e convergentes.

Teste gratuito com Bookey

Digitalize para baixar

5.5 Otimização com Grandes Conjuntos de Dados

Lidar com grandes conjuntos de dados exige estratégias de computação eficientes, como o gradiente descendente estocástico (SGD), aproveitando mini-batches para aproximar gradientes de forma econômica. Essa

Instale o app Bookey para desbloquear o texto completo e o áudio

Teste gratuito com Bookey

Ad

Por que o Bookey é um aplicativo indispensável para amantes de livros

Conteúdo de 30min
Quanto mais profunda e clara for a interpretação que fornecemos, melhor será sua compreensão de cada título.

Clipes de Ideias de 3min
Impulsione seu progresso.

Questionário
Verifique se você dominou o que acabou de aprender.

E mais
Várias fontes, Caminhos em andamento, Coleções...

Teste gratuito com Bookey

Capítulo 5 Resumo: Redes neurais e aprendizado profundo

Resumo do Capítulo: Redes Neurais e Aprendizado Profundo

Introdução às Redes Neurais e Aprendizado Profundo

- As redes neurais ampliam a regressão linear e logística empilhando múltiplos modelos para formar estruturas hierárquicas, permitindo modelar relações complexas entre entradas e saídas. O aprendizado profundo, um subconjunto do aprendizado de máquina, foca nesses modelos hierárquicos.
- O capítulo apresenta as redes neurais começando pela generalização da regressão linear para uma rede de duas camadas e, em seguida, para redes profundas. Também abrange redes específicas para imagens e técnicas de treinamento.

6.1 O Modelo de Rede Neural

- **Regressão Linear Generalizada:** Melhora a regressão linear tradicional utilizando funções de ativação não lineares (ex: logística e ReLU) para modelar relações complexas.
- **Rede Neural de Duas Camadas:** Aumenta a complexidade utilizando

Teste gratuito com Bookey

Digitalize para baixar

múltiplos modelos de regressão linear generalizada para formar camadas, agregando suas saídas como entrada para outro modelo de regressão.
- **Rede Neural Profunda:** Estende ainda mais o modelo empilhando múltiplas camadas, permitindo que ele lide com tarefas complexas como a classificação de imagens.

6.2 Treinamento de uma Rede Neural

- **Problema de Otimização:** Envolve encontrar os melhores parâmetros através de funções de custo adaptadas para tarefas de regressão ou classificação (usando funções de perda como erro quadrático e entropia cruzada).
- **Busca Baseada em Gradiente:** Atualiza parâmetros iterativamente através de técnicas como o gradiente descendente estocástico.
- **Algoritmo de Retropropagação:** Computa de forma eficiente os gradientes necessários para atualizações de parâmetros, sendo crucial para o treinamento da rede.

6.3 Redes Neurais Convolucionais (CNNs)

- Projetadas para dados de entrada em grade, principalmente imagens, as CNNs utilizam camadas convolucionais informadas pelas estruturas das imagens para melhorar a eficiência.
- **Conceitos da Camada Convolucional:**

Teste gratuito com Bookey
Digitalize para baixar

- **Interações Esparsas e Compartilhamento de Parâmetros:** Reduzem a complexidade do modelo ao focar em pequenas regiões da imagem e reutilizar parâmetros por toda a rede.
- **Stride e Pooling:** Métodos como convoluções com stride e camadas de pooling reduzem as dimensões dos dados enquanto mantêm características importantes das imagens.
- **Expansão de Canal:** Vários filtros representam diferentes aspectos da imagem, organizados como canais, produzindo representações ricas de características.

6.4 Dropout

- **Técnica de Regularização:** Treina um conjunto de sub-redes dentro de uma única rede, descartando aleatoriamente unidades e criando variações.
- **Treinamento e Predição:** Usa os gradientes das sub-redes para o treinamento, e durante a predição, a rede completa usa pesos escalonados para aproximar os resultados médios de todas as sub-redes.

6.5 Leitura Adicional

- Enfatiza o ressurgimento das redes neurais através dos avanços em hardware, computação paralela e desenvolvimento de algoritmos, especialmente em tarefas de reconhecimento de imagens.

Teste gratuito com Bookey

Digitalize para baixar

Exemplos de Aplicações e Reflexões:

- Exemplos incluem a classificação de dígitos usando conjuntos de dados MNIST, ilustrando as estruturas das redes e dinâmicas de treinamento.
- As reflexões instigam a compreensão de conceitos mais profundos, como as implicações da estrutura da rede e os efeitos das camadas de pooling.

Este capítulo fornece um guia abrangente para entender a estrutura das redes neurais, princípios de treinamento e aplicações práticas, estabelecendo a base para empreendimentos avançados em aprendizado de máquina.

Teste gratuito com Bookey

Digitalize para baixar

Capítulo 6 Resumo: Métodos de conjunto: Bagging e boosting

Capítulo 7: Métodos de Conjunto: Bagging e Boosting

Nos capítulos anteriores, discutimos diversos modelos de aprendizado de máquina, como k-NN e redes neurais profundas. Este capítulo se aprofunda nos métodos de conjunto, uma técnica na qual múltiplos modelos básicos são combinados para melhorar a previsão por meio da "sabedoria das multidões". Ao treinar cada modelo de forma diferente, suas previsões coletivas superam as individuais. As previsões de um conjunto são obtidas pela média ou votação, resultando em maior precisão quando estruturadas corretamente.

7.1 Bagging

Bagging, que significa agregação bootstrap, é uma técnica de redução de variância. Ela envolve a criação de múltiplas variações dos dados de treinamento através do bootstrapping—amostrando aleatoriamente subconjuntos sobrepostos—e treinando um modelo básico em cada um. Isso reduz a variância sem aumentar o viés, diminuindo assim o risco de overfitting. O bagging é particularmente benéfico para modelos como árvores de classificação profundas ou k-NN, que têm baixo viés, mas alta

Teste gratuito com Bookey

Digitalize para baixar

variância. Por exemplo, ao utilizar árvores de regressão, o bagging mantém o viés baixo enquanto reduz a variância, levando a predições mais estáveis. O método bootstrap envolve amostragem com reposição, permitindo que alguns pontos de dados apareçam várias vezes, enquanto outros podem não aparecer. Essa técnica ajuda a simular múltiplos conjuntos de dados a partir de uma única amostra, algo crucial quando coletar novos dados é impraticável.

A principal vantagem do bagging é a redução da variância através da média das previsões do conjunto. Na prática, mesmo com muitos membros no conjunto, o modelo não sofre com aumento do viés ou overfitting. Em vez disso, a média de previsões aleatórias, mas identicamente distribuídas, minimiza efetivamente a variância. O número de membros do conjunto (\(B\)) é frequentemente relacionado a restrições computacionais, em vez de considerações de viés ou variância.

Os modelos de bagging utilizam a estimativa de erro "out-of-bag", onde cerca de 63% dos dados participam de cada conjunto de dados bootstrapped, para estimar o erro sem precisar de validação cruzada intensiva.

7.2 Florestas Aleatórias

As florestas aleatórias, uma extensão do bagging, adicionam aleatoriedade para decorrelar modelos e reduzir ainda mais a variância. Elas são aplicáveis

Teste gratuito com Bookey

Digitalize para baixar

a modelos baseados em árvore. Cada árvore é treinada com um subconjunto de variáveis de entrada selecionadas aleatoriamente, reduzindo as correlações entre os membros. Embora isso aumente a variância de cada árvore individual, a previsão média do conjunto se beneficia da decorrelação, como demonstrado em problemas de classificação binária. As florestas aleatórias evitam o overfitting mesmo com um grande número de membros, embora as demandas computacionais se multipliquem.

O parâmetro de ajuste \(q\), que indica o número de entradas consideradas para divisões, pode ser determinado por regras práticas ou métodos de estimativa de erro. Florestas aleatórias são adequadas para implementações paralelizadas, mitigando seus custos computacionais.

7.3 Boosting e AdaBoost

Boosting, outro método de conjunto, tem como foco a redução do viés em modelos de alto viés, treinando sequencialmente modelos fracos, cada um corrigindo os erros dos predecessores. Ao contrário do bagging, o boosting constrói flexibilidade transformando modelos fracos em um conjunto forte, adequado para modelos de alto viés, como árvores de decisão rasas ou "stumps".

AdaBoost (Adaptive Boosting) é uma implementação prática de boosting, combinando múltiplos modelos fracos através de uma votação de maioria

Teste gratuito com Bookey

Digitalize para baixar

ponderada, melhorando a precisão das previsões coletivas. O treinamento sequencial enfatiza os pontos de dados difíceis, aprimorando a classificação. Compreender o AdaBoost envolve entender o ajuste estratégico dos pesos do modelo com base no desempenho de cada iteração—enfatizando dados mal classificados em rodadas subsequentes.

7.4 Gradient Boosting

Uma interpretação estatística de boosting apresenta-o como o treinamento de um modelo aditivo. Essa estrutura substitui a perda exponencial do AdaBoost por alternativas robustas, como a perda logística, mitigando a sensibilidade ao ruído. Cada etapa do boosting envolve a seleção de novos membros do conjunto com base nos gradientes dos componentes anteriores para otimizar o modelo completo de forma iterativa. A diferenciabilidade de segunda ordem da função de perda escolhida é crucial para a computação de gradientes, essencial para a metodologia de gradient boosting.

Com uma visão mais ampla do aprendizado de conjuntos, o gradient boosting atende tanto tarefas de classificação quanto de regressão, construindo sobre modelos fracos—geralmente baseados em árvore—em iterações corretivas repetitivas. Esta seção culmina apresentando o AdaBoost e o gradient boosting aplicados à classificação de músicas, demonstrando aplicações práticas e desempenhos comparativos.

Teste gratuito com Bookey

Digitalize para baixar

7.5 Leitura Adicional

Leituras chave sobre bagging e florestas aleatórias podem ser rastreadas a trabalhos fundamentais de Breiman e Ho, enquanto a evolução do boosting, popularizada por AdaBoost de Freund e Schapire, se expandiu através de contribuições significativas, incluindo o desenvolvimento de estruturas de gradient boosting por Friedman. Implementações modernas como XGBoost e LightGBM exemplificam os avanços contínuos neste campo.

Seção	Resumo
Visão Geral do Capítulo	Métodos de conjunto aprimoram a precisão das previsões ao combinar vários modelos. A média ou a votação das saídas resultam em previsões mais precisas.
7.1 Bagging	- Reduz a variância sem aumentar o viés, utilizando conjuntos de dados bootstrap. - É benéfico para modelos com alta variância (por exemplo, árvores profundas, k-NN). - A estimativa "out-of-bag" fornece uma aproximação do erro. - O tamanho do conjunto é limitado por restrições computacionais.
7.2 Florestas Aleatórias	- Expande o bagging com aleatoriedade nas características de entrada para modelos de árvore. - Reduz a correlação entre as árvores, diminuindo a variância geral. - Exige alto poder computacional, mas suporta processamento paralelo. - O parâmetro de determinação do corte (q) influencia o desempenho e pode ser ajustado.
7.3 Boosting	- Minimiza o viés ao treinar sucessivamente modelos fracos que

More Free Book

undefined

Seção	Resumo
e AdaBoost	corrigem os erros dos anteriores. - O AdaBoost ajusta os pesos para corrigir classificações incorretas, melhorando o desempenho dos modelos fracos.
7.4 Gradient Boosting	- Constrói um modelo aditivo utilizando otimização gradiente. - Utiliza funções de perda robustas para aumentar a estabilidade. - É adequado tanto para tarefas de classificação quanto de regressão. - Demonstra aplicações como a classificação de músicas.
7.5 Leituras Adicionais	- Cita pesquisas fundamentais e pioneiros como Breiman, Ho, Freund, Schapire e Friedman. - Destaca ferramentas modernas como o XGBoost e LightGBM como avanços.

More Free Book

undefined

Capítulo 7 Resumo: Transformações não lineares de entrada e núcleos

Claro! Aqui está a tradução do texto em inglês para expressões em português, com uma linguagem natural e acessível para leitores que gostam de ler livros.

Capítulo 8 deste livro explora as transformações não lineares de entrada e os métodos de núcleo em aprendizado de máquina, ampliando as ideias introduzidas no Capítulo 3. O capítulo discute principalmente como as transformações não lineares das características de entrada e o truque do núcleo podem aprimorar modelos tradicionais, como a regressão linear e k-NN, tornando-os mais flexíveis e poderosos.

Conceitos-Chave:

Transformações Não Lineares de Entrada Tradicionalmente, modelos como a regressão linear se concentram em modelar as saídas como uma combinação linear das entradas. No entanto, ao transformar as entradas de forma não linear (por exemplo, incorporando características polinomiais), conseguimos tornar esses modelos mais adaptáveis a estruturas de dados complexas. O capítulo explica que, ao transformar uma única entrada \(x \)

Teste gratuito com Bookey

Digitalize para baixar

usando uma abordagem polinomial, o modelo continua sendo uma regressão linear, onde as transformações não lineares servem como novas características de entrada.

Métodos de Núcleo: O truque do núcleo permite a implementação de modelos utilizando um número infinito de características sem computá-las explicitamente. Os núcleos calculam produtos internos em um espaço de alta dimensão de forma implícita, permitindo que os modelos capturem padrões complexos de maneira mais eficiente. Notavelmente, os métodos de núcleo são apresentados através das máquinas de vetor de suporte (SVMs), que empregam regressão de vetor de suporte e classificação de vetor de suporte, caracterizadas pelo uso de uma função de perda hinge para induzir esparsidade em suas soluções, focando principalmente nos vetores de suporte — pontos de dados chave que definem a fronteira de decisão do modelo.

Principais Componentes:

8.1 Criação de Características por Transformações Não Lineares de Entrada: Esta seção discute as sutilezas do que constitui um modelo "linear" e as possibilidades desbloqueadas ao redefinir entradas por meio de transformações não lineares. O exemplo ilustra como a adição de termos polinomiais às entradas altera a curva de resposta do modelo, sem fazer com que ele perca sua caracterização como um método linear.

Teste gratuito com Bookey

Digitalize para baixar

8.2 Regressão de Ridge com Núcleo: Estende a regressão linear através do truque do núcleo para lidar com infinitas características. A regularização torna-se crucial para evitar o sobreajuste ao adicionar complexidade ao modelo. Esta seção detalha o processo de uso de núcleos como o polinomial ou o gaussiano, destacando como eles influenciam a flexibilidade do modelo.

8.3 Regressão de Vetor de Suporte: Apresenta uma alternativa de regressão baseada em núcleo que utiliza a função de perda \(\epsilon\)-insensível, levando a soluções esparsas que focam apenas em pontos de dados críticos (vetores de suporte). Esta propriedade economiza computação durante as fases de predição, já que apenas um subconjunto dos dados de treinamento influencia as fronteiras de decisão do modelo.

8.4 Teoria do Núcleo: Fornece uma base teórica para os métodos de núcleo, explorando a conexão entre núcleos e mapeamentos de características, e o conceito de espaço de Hilbert reproduzível. Também oferece conselhos práticos sobre como escolher núcleos e discute a formação de núcleos válidos, enfatizando os núcleos semidefinidos positivos devido às suas propriedades matemáticas garantidas.

8.5 Classificação de Vetores de Suporte: Semelhante à regressão de vetor de suporte, mas aplicada a problemas de classificação. Utiliza uma função

Teste gratuito com Bookey

Digitalize para baixar

de perda hinge para garantir que apenas os pontos de dados próximos à margem de decisão (vetores de suporte) influenciem a fronteira de classificação. Esta seção discute também a formulação matemática necessária para implementar classificadores de vetor de suporte usando diferentes tipos de núcleos, como o linear ou o exponencial quadrático.

Conclusão:

O capítulo enfatiza que, embora os métodos de núcleo elevem a flexibilidade do modelo e capturem relações não lineares nos dados de forma eficaz, eles exigem uma escolha cuidadosa do núcleo e ajuste de parâmetros. Através dos núcleos, mesmo modelos simples como a regressão linear ou k-NN podem se adaptar a cenários de dados complexos do mundo real, tornando-os ferramentas poderosas para engenheiros e cientistas que enfrentam desafios em aprendizado de máquina.

Apêndice - Teorema do Representante e Derivação da Classificação de Vetores de Suporte:

O teorema do representante afirma que a solução de modelos baseados em núcleo regularizados pode ser expressa em termos de dados de treinamento, destacando a elegância das formulações duais empregadas nos métodos de vetor de suporte. O apêndice também aborda a derivação matemática necessária para a transição de formas primais para duais, uma parte integral

Teste gratuito com Bookey

Digitalize para baixar

da implementação de classificadores de vetor de suporte na prática.

Espero que a tradução atenda às suas expectativas! Se precisar de mais ajuda, é só avisar.

Teste gratuito com Bookey

Digitalize para baixar

Capítulo 8: A abordagem bayesiana e os processos gaussianos

Capítulo 9: A Abordagem Bayesiana e Processos Gaussianos

Este capítulo explora a abordagem bayesiana em aprendizado de máquina, contrastando-a com o método paramétrico que busca um único valor de parâmetro ideal para ajustar modelos aos dados. Aqui, o aprendizado é reimaginado como a determinação de uma distribuição de possíveis valores de parâmetros condicionados aos dados observados. Essa perspectiva probabilística nos permite prever não apenas um único resultado, mas uma gama de resultados possíveis, cada um com sua probabilidade. A estrutura bayesiana é distinta tanto teoricamente quanto filosoficamente, oferecendo uma metodologia robusta aplicável a uma ampla gama de cenários práticos em aprendizado de máquina supervisionado.

9.1 A Ideia Bayesiana

Na abordagem bayesiana, os parâmetros do modelo são tratados como variáveis aleatórias. O foco está em formar uma distribuição conjunta sobre todas as saídas e parâmetros, Ø5Ü](y " Ø5ß= | X). Na p
de estimar um parâmetro pontual para calcular a distribuição posterior dos parâmetros Ø5Ü](Ø5ß= | y), realizada por meio de leis

Teste gratuito com Bookey

Digitalize para baixar

teorema de Bayes. O teorema, Ø5Ü](Ø5ß= | y) = Ø5Ü](y |

- **Verossimilhança** Ø5Ü](y | Ø5ß=): distribuição dos dados dado os parâmetros.
- **Prior** Ø5Ü](Ø5ß=): distribuição dos parâmetros antes da observação dos dados.
- **Posterior** Ø5Ü](Ø5ß= | y): distribuição atualizada dos parâmetros após a observação dos dados.

As distribuições de probabilidade neste contexto representam crenças sobre parâmetros ou previsões em condições de incerteza. As previsões de um modelo bayesiano, a robustez do modelo contra o sobreajuste, especialmente com dados limitados, e a atualização sequencial de crenças com novos dados tornam essa abordagem atraente.

9.2 Regressão Linear Bayesiana

A estrutura bayesiana aplicada à regressão linear ilustra a transição de uma estimativa pontual para uma inferência probabilística. Uma distribuição gaussiana multivariada desempenha um papel crucial. A regressão linear bayesiana envolve:

- Definir distribuições anteriores para os parâmetros do modelo.
- Usar dados observados para atualizar a anterior em uma distribuição posterior Ø5Ü](Ø5ß= | y).
- Derivar previsões por meio de distribuições preditivas posteriores.

Teste gratuito com Bookey

Digitalize para baixar

Aqui, os conceitos de regularização alinham-se naturalmente com os princípios bayesianos, oferecendo insights sobre sua utilidade na mitigação do sobreajuste.

9.3 O Processo Gaussiano

O Processo Gaussiano (PG) estende a abordagem bayesiana a modelos não paramétricos, considerando funções inteiras como entidades estocásticas para inferir previsões. O PG é um processo estocástico contínuo visto através de uma lente gaussiana multivariada para qualquer conjunto finito de entradas. Sua função de covariância (kernel) desempenha um papel essencial na definição de correlações entre os valores da função. O PG permite a regressão com quantificação da incerteza para previsões, mostrando flexibilidade através de kernels que especificam características inerentes da função como suavidade.

9.4 Aspectos Práticos do Processo Gaussiano

Escolher o kernel adequado e ajustar hiperparâmetros (com métodos como maximização da verossimilhança marginal) são fundamentais para implementar Processos Gaussianos de maneira eficaz. Essas escolhas impactam significativamente o desempenho preditivo e a eficiência computacional.

Teste gratuito com Bookey

Digitalize para baixar

9.5 Outros Métodos Bayesianos em Aprendizado de Máquina

O paradigma bayesiano abrange muito mais do que regressão, englobando uma variedade de modelos que requerem técnicas computacionais

Instale o app Bookey para desbloquear o texto completo e o áudio

Teste gratuito com Bookey

Ad

App Store
Escolha dos Editores

22k avaliações de 5 estrelas

Feedback Positivo

Afonso Silva

s cada resumo de livro não só
o, mas também tornam o
m divertido e envolvente. O
tou a leitura para mim.

Fantástico! Brígida Santos

Estou maravilhado com a variedade de livros e idiomas que o Bookey suporta. Não é apenas um aplicativo, é um portal para o conhecimento global. Além disso, ganhar pontos para caridade é um grande bônus!

Fi

O
só
o
O

na Oliveira

correr as
ém me dá
omprar a
ar!

Adoro! Duarte Costa

Usar o Bookey ajudou-me a cultivar um hábito de leitura sem sobrecarregar minha agenda. O design do aplicativo e suas funcionalidades são amigáveis, tornando o crescimento intelectual acessível a todos.

Economiza tempo!

O Bookey é o meu apli
crescimento intelectua
perspicazes e lindame
um mundo de conheci

Aplicativo incrível! Estevão Pereira

Eu amo audiolivros, mas nem sempre tenho tempo para ouvir o livro inteiro! O Bookey permite-me obter um resumo dos destaques do livro que me interessa!!! Que ótimo conceito!!! Altamente recomendado!

Aplicativo lindo

Este aplicativo é um salva-vidas para
de livros com agendas lotadas. Os re
precisos, e os mapas mentais ajudar
o que aprendi. Altamente recomend

Teste gratuito com Bookey

Certainly! Here's the translation of "Chapter 9" into Portuguese:

Capítulo 9

If you need anything else or more text to be translated, feel free to ask! Resumo: Modelos generativos e aprendizado a partir de dados não rotulados.

Capítulo 10: Modelos Geradores e Aprendizado com Dados Não Rotulados

Visão Geral

Este capítulo faz a transição dos modelos discriminativos, que predizem saídas com base em entradas fornecidas, para os modelos geradores. Ao contrário dos modelos discriminativos, que apenas descrevem a distribuição condicional de uma saída dada uma entrada, os modelos geradores descrevem a distribuição conjunta de entradas e saídas. Isso proporciona uma compreensão mais rica dos dados e permite a geração de dados sintéticos ou a identificação de padrões de dados sem a necessidade de saídas rotuladas.

Introdução aos Modelos Geradores

Os modelos geradores descrevem como os dados de entrada e saída são

Teste gratuito com Bookey

Digitalize para baixar

gerados em conjunto. Um modelo notável é o Modelo de Mistura Gaussiana (GMM), que representa os dados como misturas de distribuições gaussianas, servindo para análises discriminantes, aprendizado semi-supervisionado e não supervisionado, como agrupamento.

Modelo de Mistura Gaussiana (GMM) e Análise Discriminante

- **Conceitos Básicos do GMM**: Um GMM descreve um modelo probabilístico no qual as variáveis de entrada seguem uma distribuição gaussiana dentro de cada classe. Ele utiliza a probabilidade conjunta \(p(x, y) = p(x | y) p(y)\), onde \(y\) é uma categoria com uma distribuição gaussiana de \(x\) para cada classe.
- **Aprendizado de Parâmetros**: Em um cenário de aprendizado supervisionado, os parâmetros do GMM são aprendidos maximizando a verossimilhança logarítmica dos dados observados. Isso resulta em métodos como Análise Discriminante Linear (LDA) e Análise Discriminante Quadrática (QDA) para classificação.
- **Aplicações Semi-supervisionadas e de Agrupamento**: Os GMMs permitem o agrupamento de dados não rotulados e o aprendizado a partir de dados parcialmente rotulados. A imputação de rótulos e o algoritmo de Expectativa-Maximização (EM) desempenham papéis cruciais na abordagem de desafios semi-supervisionados, atualizando iterativamente os parâmetros do modelo com base nas previsões de dados não rotulados.

Aprendizado Não Supervisionado e Análise de Agrupamento

Teste gratuito com Bookey

Digitalize para baixar

- **Abordando Cenários Não Supervisionados**: Os GMMs são adaptados para aprendizado não supervisionado ao considerar que os dados contêm variáveis latentes de cluster. Isso envolve aprender um modelo conjunto \(p(x, y)\) apenas a partir das entradas para identificar clusters inerentes aos dados.
- **Visão Geral do Algoritmo EM**: O algoritmo de Expectativa-Maximização (EM) é empregado para refinar iterativamente os parâmetros do modelo, alternando entre a estimativa de variáveis ocultas e a otimização de parâmetros.
- **Agrupamento Ø5ÜX-means**: Como um método alt agrupamento, o Ø5ÜX-means particiona os dados em cl variância intracluster. Oferece uma estrutura mais simples, porém menos flexível em comparação ao agrupamento probabilístico do GMM.

Modelos Geradores Avançados
- **Limitações e Extensões Além da Suposição Gaussiana**: Embora os GMMs assumam distribuições gaussianas, modelos geradores profundos oferecem maior flexibilidade usando transformações complexas de distribuições mais simples, como no caso das redes neurais.
- **Fluxos Normalizantes**: Essa técnica modela os dados como transformações de distribuições mais simples e envolve mapeamentos reversíveis com jacobianos computacionalmente viáveis para o aprendizado de parâmetros.
- **Redes Geradoras Adversariais (GANs)**: As GANs aprendem sem

Teste gratuito com Bookey
Digitalize para baixar

verossimilhanças explícitas, gerando dados por meio de treinamento adversarial, onde uma rede geradora cria dados e uma rede discriminadora avalia sua autenticidade.

Aprendizado de Representação e Redução de Dimensionalidade
- **Autoencoders**: Esses aprendem representações de baixa dimensão impondo um gargalo que comprime os dados de entrada, reconstruindo as entradas originais da forma mais próxima possível.
- **Análise de Componentes Principais (PCA)**: A PCA, vista como uma solução de autoencoder linear, encontra uma representação de dimensionalidade reduzida por meio de transformações ortogonais que maximizam a variância dos dados.

Considerações Finais
Os modelos geradores oferecem estruturas ricas para compreender e gerar dados sintetizados. Embora envolvam mais suposições do que os modelos discriminativos, eles desbloqueiam possibilidades em contextos onde a rotulagem é cara ou impraticável, fazendo a ponte entre paradigmas de aprendizado supervisionado e não supervisionado.

Sugestões de Leitura Adicional
Consulte discussões mais detalhadas em trabalhos de Bishop (2006), Hastie et al. (2009) e Goodfellow et al. (2016), além das obras fundamentais sobre GANs e fluxos normalizantes (Goodfellow et al., 2014; Kobyzev et al.,

Teste gratuito com Bookey
Digitalize para baixar

2020). Para aplicações e expansões, veja a literatura sobre autoencoders variacionais e métodos de aprendizado semi-supervisionado.

Teste gratuito com Bookey

Digitalize para baixar

Capítulo 10 Resumo: Aspectos do usuário em aprendizado de máquina

Capítulo 11 do livro concentra-se nos aspectos práticos do usuário em aprendizado de máquina, especialmente no contexto do aprendizado supervisionado. Este resumo delineia os pontos-chave no desenvolvimento e aprimoramento de modelos de aprendizado de máquina, garantindo o uso ideal de recursos limitados, como dados e tempo.

Destaques do Capítulo:

1. **Definindo o Problema de Aprendizado de Máquina:**

- O desenvolvimento de aprendizado de máquina é um processo iterativo—treinar, avaliar e melhorar. O objetivo é avaliar melhorias de forma eficiente sem implantações frequentes em ambientes de produção.
- A estratégia envolve o uso de uma divisão de dados bem definida: dados de treinamento para aprender o modelo, dados de validação para ajustar hiperparâmetros e dados de teste para avaliação final.
- Quando os dados são limitados, técnicas como validação cruzada k-fold podem complementar métodos de validação hold-out. É crucial garantir que os conjuntos de dados de validação e teste representem a distribuição esperada em produção para evitar "mirar no alvo errado".

Teste gratuito com Bookey

Digitalize para baixar

- O risco de overfitting aos dados de validação pode ser minimizado expandindo o tamanho do conjunto de dados de validação e fazendo uma partição cuidadosa para prevenir vazamentos de grupo, especialmente em domínios como imagem médica.

2. **Aprimorando um Modelo de Aprendizado de Máquina:**

- O processo de aprimoramento do modelo é iterativo e pode envolver a tentativa de diferentes modelos ou ajustes de hiperparâmetros.
- Comece com abordagens simples e avance gradualmente para modelos mais complexos. A depuração e verificação da correção do modelo devem preceder melhorias elaboradas.
- O capítulo enfatiza a manutenção de um equilíbrio entre a minimização do erro de treinamento e a redução da lacuna de generalização (a diferença entre o desempenho em treinamento e validação).
- A avaliação de curvas de aprendizado e a utilização de análise de erros ajudam a priorizar esforços de coleta de dados e identificar áreas com potencial para melhorias no modelo.

3. **Estratégias quando os Dados são Limitados:**

- Em situações com dados limitados, várias estratégias podem ajudar:
- Aumentar os dados de treinamento com pequenas modificações ou aproveitar conjuntos de dados não relacionados que ainda compartilhem

Teste gratuito com Bookey

Digitalize para baixar

semelhanças.

- O uso de aprendizado por transferência para transferir conhecimento de modelos treinados em grandes conjuntos de dados para novas tarefas relacionadas é destacado.

- Utilizar aprendizado auto-supervisionado ou semi-supervisionado para obter insights de conjuntos de dados não rotulados.

4. **Addressing Practical Data Issues:**

- Desafios práticos como outliers, dados faltantes e características redundantes devem ser abordados para garantir a robustez do modelo.

- Os outliers exigem consideração cuidadosa—decidir se são ruído ou fenômenos de interesse.

5. **Confiabilidade dos Modelos de Aprendizado de Máquina:**

- Compreender e confiar em modelos de aprendizado de máquina supervisionados pode ser desafiador, especialmente para modelos complexos como o aprendizado profundo.

- Embora modelos mais simples ofereçam vantagens de interpretabilidade, eles podem carecer do poder preditivo de sistemas mais complexos. Esforços estão em andamento na área para melhorar a interpretabilidade e a robustez contra exemplos adversários e cenários de pior caso.

Teste gratuito com Bookey

Digitalize para baixar

6. **Exploração Adicional:**

- O capítulo menciona trabalhos de Ng (2019) e Burkov (2020) para mais insights sobre aspectos do usuário em aprendizado de máquina, bem como literatura sobre técnicas de aumento de dados e interpretação de modelos.

O capítulo enfatiza uma abordagem estruturada e informada para o desenvolvimento de soluções de aprendizado de máquina, focando em considerações práticas e melhorias iterativas para lidar efetivamente com desafios do mundo real.

Teste gratuito com Bookey

Digitalize para baixar

Pensamento Crítico

Ponto Chave: Aprendizado por Transferência

Interpretação Crítica: Imagine aproveitar o poder de uma vasta quantidade de informações disponíveis através de modelos treinados em outros domínios e aplicá-las às suas necessidades específicas. O aprendizado por transferência permite que você faça exatamente isso. Ao utilizar modelos pré-existentes treinados com grandes e diversos conjuntos de dados, você abre as portas para possibilidades infinitas, mesmo quando começa com dados limitados. Pense nisso como estar sobre os ombros de gigantes, utilizando a sabedoria coletiva codificada em modelos sofisticados para resolver seus problemas únicos. Isso é mais do que um atalho técnico; é uma mudança de paradigma que acelera seu crescimento e fortalece você, mesmo em ambientes com recursos restritos. Adote essa estratégia para desbloquear um potencial que você nunca soube que existia, alcançando objetivos com criatividade e inovação.

Teste gratuito com Bookey

Digitalize para baixar

Capítulo 11 Resumo: Ética na aprendizagem de máquina

Resumo do Capítulo 12: Ética em Aprendizado de Máquina

No Capítulo 12 de "Aprendizado de Máquina: Um Primeiro Curso para Engenheiros e Cientistas", de David Sumpter, o foco está nos desafios éticos que surgem nas aplicações de aprendizado de máquina. Estas questões muitas vezes decorrem de escolhas de design aparentemente neutras que levam a consequências inesperadas para os usuários e para a sociedade. O capítulo defende uma abordagem de ética através da conscientização, enfatizando a importância de entender esses impactos éticos em vez de se apoiar apenas em soluções técnicas.

12.1 Justiça e Funções de Erro

A primeira seção aborda a justiça no aprendizado de máquina, começando pela escolha de uma função de erro. Sumpter ilustra isso com um exemplo fictício de um modelo de recomendação de cursos universitários para candidatos suecos e não suecos. Apesar de um desempenho geral igual, as taxas de falsos negativos revelam uma disparidade, já que candidatos não suecos interessados têm mais chances de serem recomendados incorretamente do que suecos. Isso ressalta que a justiça depende do contexto — o que é justo em um cenário pode ser injusto em outro. O

Teste gratuito com Bookey

Digitalize para baixar

algoritmo Compas, usado na sentença criminal, exemplifica a impossibilidade matemática de alcançar justiça em todos os critérios desejáveis simultaneamente. Essa seção enfatiza que a justiça é subjetiva e exige conscientização sobre decisões de design e suas implicações.

12.2 Afirmativas Enganosas Sobre o Desempenho

O crescimento rápido do aprendizado de máquina levou a afirmações exageradas de desempenho, influenciadas por interesses comerciais. Por exemplo, enquanto os feitos de aprendizado por reforço da Google DeepMind em jogos foram aclamados como marcos, suas capacidades foram, por vezes, superestimadas. Um exemplo crítico é o algoritmo Compas, usado na sentença criminal. Embora tenha demonstrado alto desempenho em benchmarks técnicos como AUC, sua eficácia no mundo real era comparável a previsões casuais de humanos e acrescentava pouco ao conhecimento existente sobre a previsão de crimes.

O capítulo também critica as afirmações exageradas de Elon Musk sobre as capacidades de IA da Tesla e a precisão dos modelos de aprendizado de máquina usados na previsão de personalidade pela Cambridge Analytica. Os perigos das alegações hiperbólicas são agravados por terminologia complexa, que pode obscurecer as implicações reais para não especialistas. O capítulo encoraja o uso de linguagem acessível para transmitir o desempenho e as limitações dos modelos, promovendo uma comunicação

Teste gratuito com Bookey

Digitalize para baixar

honesta com as partes interessadas.

12.3 Limitações dos Dados de Treinamento

Sumpter introduz o conceito de modelos de aprendizado de máquina como "papagaios estocásticos", refletindo a ideia de que os modelos apenas repetem padrões vistos nos dados de treinamento sem compreender. Esta analogia ajuda a explicar por que os modelos falham quando enfrentam variações não presentes em seus dados de treinamento. O capítulo discute os preconceitos inerentes aos dados de treinamento, como disparidades raciais e de gênero em sistemas de reconhecimento facial, que perpetuam preconceitos sociais existentes.

Quando aplicados a grandes conjuntos de dados não verificados, os modelos de linguagem podem, inadvertidamente, produzir saídas tendenciosas ou incorretas, como repetir teorias da conspiração. Isso ressalta a necessidade de conjuntos de dados eticamente obtidos e equilibrados e destaca os riscos de assumir que os modelos entendem os dados em que foram treinados. Sumpter enfatiza a importância de reconhecer a influência de fatores históricos e sociais nos dados e destaca que a neutralidade do modelo é um mito.

Conclusão

Teste gratuito com Bookey

Digitalize para baixar

De modo geral, o Capítulo 12 do livro de Sumpter enfatiza a importância da conscientização ao enfrentar desafios éticos no aprendizado de máquina. Embora estar ciente dessas questões não as resolva, é um ponto de partida crucial. O capítulo encoraja engenheiros e cientistas a se comunicarem de forma clara, a reconhecerem preconceitos e a considerarem ativamente as implicações éticas em seu trabalho. Leituras adicionais, incluindo as de Bender et al., David Sumpter e Cathy O'Neill, são recomendadas para uma compreensão mais aprofundada dessas questões complexas.

Teste gratuito com Bookey

Digitalize para baixar

Pensamento Crítico

Ponto Chave: Abordagem Ética-Através-da-Conscientização

Interpretação Crítica: Imagine ter em mãos o poder do aprendizado de máquina, uma ferramenta capaz de causar um impacto monumental, mas repleta de dilemas éticos ocultos. Adote a abordagem ética através da conscientização para navegar por essas águas turvas. Este capítulo enfatiza a importância de compreender as ondulações éticas que suas escolhas de design provocam na sociedade. Em vez de confiar cegamente na fachada da neutralidade da tecnologia, reconheça que a intenção, o design e a aplicação dos algoritmos de aprendizado de máquina exigem uma avaliação consciente de justiça e impacto. Ao embarcar nessa jornada de conscientização, você revela a natureza subjetiva da justiça, garantindo que suas contribuições atendam não apenas ao sucesso técnico imediato, mas também ressoem uma harmonia ética mais ampla. Este princípio pode revolucionar a forma como você percebe a responsabilidade, promovendo um mundo onde a inovação é orientada igualmente pela empatia e compreensão.

Teste gratuito com Bookey

Digitalize para baixar