ANGÉLICA DE SOUSA PINTO

UM SISTEMA PARA PREDIÇÃO DE TEMPO DE PERCURSO DE ROTA

Proposta de Dissertação apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do título de Mestre em Engenharia de Produção e Sistemas Computacionais. Área de Concentração: Sistemas Computacionais.

Orientador: Prof. Dra. Flávia Bernardini

Coorientador: Prof. Dr. Patrick Barbosa Moratori

RIO DAS OSTRAS, RJ

2016

**RESUMO**

O encontro de um engarrafamento na saída e na chegada de uma cidade quando o interesse é o deslocamento causando perda de tempo no trânsito, aumentando o estress, entre outros problemas. Conseguir prever o tempo que se gasta em uma viagem, em diferentes dias da semana, no feriado, véspera de feriado, calculando o tempo necessário gastando o mínimo de tempo possível, aumenta a qualidade de vida tornando o tempo mais útil. As tecnologias existentes não apresentam alguma funcionalidade que faça a previsão do tempo gasto em uma viagem. Com a computação mais especificamente algoritmos online, é possível através de uma base de dados recolhida de vários pontos da estrada, fazer uma predição do tempo gasto de uma viagem. Os algoritmos de regressão online são aplicados no presente problema, devido dados de treinamento estejam disponíveis na memória. No entanto, o aumento exponencial dos dados armazenados na Base de Dados. O fluxo de dados é gerado em alta velocidade e o conceito a ser aprendido pode mudar ao longo do tempo. Os algoritmos aplicados neste trabalho são RNA (perceptron) online e Regressão online na base de dados, o resultado será avaliado qual o mais adequado para o problema mencionado.

**SUMÁRIO**

[**1** **INTRODUÇÃO** 2](#_Toc456992487)

[1.1TECNOLOGIAS EXISTENTES 2](#_Toc456992488)

[1.2TRABALHOS CORRELACIONADOS 2](#_Toc456992489)

[1.3OBJETIVOS 2](#_Toc456992490)

[1.4ORGANIZAÇÃO DA PROPOSTA 2](#_Toc456992491)

[**2** **ALGORITMOS DE APRENDIZADO DE MÁQUINA ONLINE PARA PROBLEMAS DE REGRESSÃO** 2](#_Toc456992492)

[2.1REGRESSÃO 2](#_Toc456992493)

[2.2 ALGORITMOS ONLINE 2](#_Toc456992494)

[2.3 REGRESSÃO LINEAR 2](#_Toc456992495)

[2.3.1FUNÇÃO DE PERDA INSENSÍVEL 2](#_Toc456992496)

[2.4REGRESSÃO ONLINE 2](#_Toc456992497)

[2.4.1 PERCEPTRON DE Ε-RAIO FIXO 2](#_Toc456992498)

[2.4.1.1PROVA DE CONVERGÊNCIA 2](#_Toc456992499)

[2.4.2PERCEPTRON DE Ρ-RAIO FIXO 2](#_Toc456992500)

[2.5RNA (PERCEPTRON) ONLINE 2](#_Toc456992501)

[2.5.1 SINAPSE 2](#_Toc456992502)

[2.5.2 O NEURÔNIO ARTIFICIAL 2](#_Toc456992503)

[2.6 SUPORTE VETOR MACHINE ONLINE 2](#_Toc456992504)

[2.6.1 SUPORTE VETOR DE REGRESSÃO 2](#_Toc456992505)

[**3** **METODOLOGIA DA PESQUISA** 2](#_Toc456992506)

[**4 ATIVIDADES E CRONOGRAMA 2**](#_Toc456992507)

[**5 Considerações finais 2**](#_Toc456992507)**5**

[**REFERÊNCIAS 2**](#_Toc456992508)**6**

1. **INTRODUÇÃO**

Um sistema de unidades móveis está instalado em vários veículos no tráfego. Estas unidades móveis incluem os dispositivos de comunicações sem fios e aparelhos que determina a localização de cada veículo. Monitorar a posição de um veículo, tal como uma função do tempo revela também a velocidade do veículo. Posição e informações de velocidade são transmitidas periodicamente pelos veículos a uma estação de monitoramento central e aos veículos vizinhos. Na estação de monitoramento central, a entrada coletiva de um conjunto de veículos é processada para fornecer um gráfico instantâneo das condições de tráfego na área. Avisos de atrasos ou atualizações sobre as condições de tráfego na estrada à frente são usados ​​como parte de um sistema de estrada para automóveis inteligentes. Veículos dentro de uma região se comunicam com outros dentro de uma rede na qual as informações de difusão são processadas localmente nos respectivos veículos para estimar possíveis problemas com antecedência e considerar o cálculo de uma estrada alternativa e/ou verificar com a estação central de monitoramento para mais informações. Se fora do alcance da estação central de monitoramento, os veículos formam uma rede de área local para a troca e atualização de informações, e quando qualquer veículo da rede está dentro do alcance da estação central de monitoramento, os dados de rede de área local são enviados para ajudar a atualizar as informações de tráfego em geral.

# OBJETIVOS

O objetivo deste trabalho é investigar o uso de algoritmos de aprendizado de máquina online para problemas de regressão, predizer o horário de partida de maneira a minimizar o tempo de viagem, por exemplo, deseja-se obter o melhor horário de partida saindo de Rio das Ostras para o Rio de janeiro, passando o menor tempo possível viajando. Os algoritmos a serem investigados são Regressão Linear online e rede neural artificial (perceptron) online.

# ORGANIZAÇÃO DA PROPOSTA

No Capítulo 2 são descritos algoritmos de problemas de regressão online a serem utilizados neste trabalho.

No Capítulo 3 é apresentada a metodologia de investigação a ser utilizada para o desenvolvimento da dissertação.

No Capítulo 4 são apresentadas as atividades a serem desenvolvidas e um cronograma de trabalho.

No Capítulo 5 são apresentadas as considerações finais.

1. **ALGORITMOS DE APRENDIZADO DE MÁQUINA ONLINE PARA PROBLEMAS DE REGRESSÃO**

A Aprendizagem de Máquina é a ciência que faz com que os computadores exerçam seu papel de forma natural sem que pareçam explicitamente programados para tal (NG, A., 2016). Na última década, a aprendizagem de máquina foi responsável pelo surgimento dos carros automáticos, recursos de reconhecimento de fala, otimizou as buscas na web e possibilitou um avanço enorme na compreensão do genoma humano estando difundida com frequência na vida das pessoas.

A forma de organizar os algoritmos de aprendizado de máquina é útil porque força a pensar sobre os papéis dos dados de entrada e o processo de preparação do modelo e seleciona o que é o mais adequado para o problema, a fim de obter o melhor resultado.

Uma característica comum em sistemas de aprendizado de máquina é através de um grande volume de informações, identificar padrões para extrair conclusões e realizar uma tarefa e quanto mais tempo e informações absorvidas, melhores se tornam as respostas.

Nas últimas décadas, com a crescente complexidade dos problemas a serem tratados computacionalmente e do volume de dados gerados por diferentes setores, tornou-se clara a necessidade de ferramentas computacionais mais sofisticadas, que fossem mais autônomas, reduzindo a necessidade de intervenção humana e dependência de especialistas. Para isso, essas técnicas deveriam ser capazes de criar por si próprias, a partir da experiência passada, uma hipótese, ou função, capaz de resolver o problema que deseja tratar. Um exemplo simples é a descoberta de uma hipótese na forma de uma regra ou conjunto de regras para definir que clientes de um supermercado devem receber material de propaganda de um novo produto, utilizando para isso dados de compras passados dos clientes cadastrados na base de dados do supermercado. A esse processo de indução de uma hipótese (ou aproximação da função) a partir da experiência passada dá-se o nome Aprendizado de Máquina (FACELI *et al*., 2011).

# REGRESSÃO e aprendizado online

O modelo de Regressão pode conter uma ou dezenas de variáveis de entrada e uma variável de saída e têm o mesmo padrão geral. Há algumas variáveis independentes que, quando tomadas em conjunto produzem um resultado, uma variável dependente. O modelo de regressão é usado para prever o resultado de uma variável dependente desconhecida, dados os valores das variáveis independentes.

Algoritmos de aprendizado online são aqueles que recebem suas entradas como uma sequência de partes e processam cada uma recebida sem conhecimento daquelas que estão por vir (FIAT, WOEGINGER, 1998). Ao processar uma parte, o algoritmo deve tomar decisões irrevogáveis que afetem sua solução. Por isso, algoritmos online são interessantes para tratar problemas em que não se tem todas as informações. Nesta seção são descritos três algoritmos de aprendizado online: regressão online, redes neurais artificiais online e máquinas de vetor suporte para regressão online.

# REGRESSÃO LINEAR online

**Regressão Linear Clássica:** Considere o problema de aproximar o conjunto de dados por uma função linear , já que a função de regressão ótima, dada pelo mínimo do funcional, + ), onde C é um valor  pré-determinado, e , são variáveis de folga representando restrições superiores e inferiores nas saídas do sistema, é uma função é desconhecida. Usando uma função de perda insensível de outra forma , a solução é dada por, , definida na Eq 1, ou alternativamente, dado pela Eq. 2, para a qual devem ser respeitadas as restrições definidas pelas Eq. 3 e 4. Resolvendo a equação 5.5 com restrições da Equação 5.7 determina os multiplicadores de Lagrange, α, α\*, e a função de regressão é dada pela Equação 5.2, onde...

|  |  |
| --- | --- |
|  | (1) |

|  |  |
| --- | --- |
|  | (2) |

|  |  |
| --- | --- |
|  | (3) |
|  | (4) |

**Regressão Online:** No contexto de aprendizado online a função candidata (geralmente chamada de hipótese) é construída através da minimização do risco empírico examinando um exemplo de treinamento por vez. Dessa maneira, inicia-se com uma hipótese inicial e, a cada iteração, o algoritmo examina um exemplo e atualiza a hipótese atual de acordo com uma regra de correção específica. Com o objetivo de derivar essa regra de correção segue-se as ideias do algoritmo Perceptron (ROSENBLATT, 1958) usando a abordagem da descida do gradiente estocástica. Considerando o risco empírico definido na seção 2, define-se o seguinte custo:

que deve ser minimizado em relação a. Assim, para cada par de pontos , a seguinte regra de correção é aplicada à hipótese atual

(2.3.1.1)

em que η > 0 é geralmente chamado de taxa de aprendizado e denota o gradiente da função de perda em relação a .

Um aspecto importante dessa abordagem é que se , o que é verdadeiro para a maioria das funções de perda, a atualização acima precisa ser efetuada somente nos casos em que. Caso contrário, a hipótese atual já atingiu o mínimo para o exemplo e não é necessário proceder qualquer correção, i.e.,. Nesse sentido, funções de perda que são baseadas no conceito de tubo são bem adequadas para esse esquema, uma vez que o exemplo somente afetará a hipótese atual caso encontre-se fora do tubo.

# Rede Neural Artificial ONLINE

**Redes Neurais Clássicas:** As redes neurais artificiais têm sua inspiração biológica originada de um neurônio típico, formado por dendritos, o corpo celular e os axônios, que junto com outros neurônios podem ter várias combinações com as mais variadas topologias. A ideia básica é sumarizada na abordagem Reducionista: reproduzindo-se com suficiente detalhe a suposta "máquina" biológica responsável pelo comportamento inteligente, (ou seja, o cérebro), um comportamento inteligente emergirá do sistema. Historicamente, os primeiros trabalhos nessa área iniciaram-se na década de 40 com a Cibernética. Todavia, somente a partir da década de 80 esta tecnologia tomou impulso para a representação do conhecimento e para o aprendizado de máquina. REFERENCIAR A FIGURA 1 A seguir são descritos os componentes do neurônio artificial.

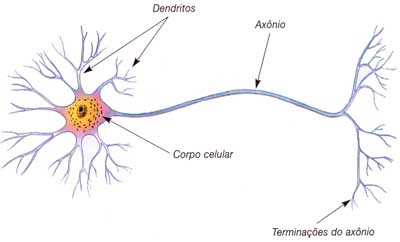


Figura 1: Esquema simplificado de um neurônio.

Fonte:(DIA A DIA EDUCAÇÃO, 2016).

**Sinapse:** No interior da membrana há uma alta concentração de moléculas específicas de neurotransmissores. Os neurotransmissores ficam armazenados em vesículas especiais denominadas vesículas sinápticas. Quando a membrana é despolarizada, isto é, quando o potencial do citoplasma torna-se maior que o potencial de repouso, os canais da membrana permitem íons como o cálcio fluir para dentro da célula, estes ativarão a liberação dos neuro-transmissores (tal como acetilcolina) constituindo-se assim a sinapse, que é a ligação entre a terminação axônica e os dendritos e que permite a propagação dos impulsos nervosos de uma célula a outra. As sinapses podem ser excitatórias ou inibitórias. As sinapses excitatórias cujos neuro-excitadores são os íons sódio permitem a passagem da informação entre neurônios e as sinapses inibitórias, cujos neuro-bloqueadores são os íons potássio, bloqueiam a atividade da célula, impedindo ou dificultando a passagem da informação.

**O Neurônio Artificial:** O corpo faz a soma ponderada do produto dos pesos, são a intensidade da força sináptica e podem ser fixos ou treináveis implementando as ligações entre as unidades e a intensidade com que o sinal é transmitido de um neurônio ao outro. O modelo do neurônio artificial de acordo com (Dayhoff, 1992) é apresentado na Figura 2. São elementos do neurônio artificial a função soma, a função de transferência e a função de ativação:

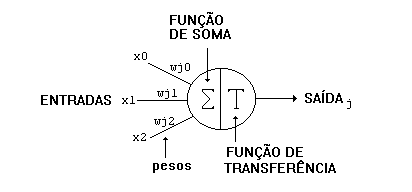


Figura 2: Modelo do Neurônio Artificial.  
Fonte: (TAFNER, 1998).

* **Função Soma** - As entradas depois de multiplicadas pelos pesos, recebem o nome de entradas ponderadas, ou unidades sigma.
* **Função de Transferência** - depois de acumulado o valor somado, o neurônio deve comparar esse valor com um valor de limiar (um valor estipulado), e atingindo-o, o valor é então passado adiante através da saída. A esse processo chamamos de função de transferência. Se o valor não atinge o limiar, o sinal não é passado adiante. Alógica neural expõe, que a intensidade dos sinais de entrada, dispara ou não, o sinal do neurônio, fazendo com que este estimule o neurônio seguinte. A função de transferência define e envia para fora do neurônio o valor passado pela função de ativação.
* **Função de ativação** - A função de ativação antecede a função de transferência, e tem por atribuição repassar o sinal para a saída do neurônio. A função de ativação é uma função de ordem interna, cuja atribuição é fazer acontecer um nível de ativação dentro do próprio neurônio, ou seja uma decisão é tomada pelo próprio neurônio sobre o que fazer com o valor resultante do somatório das entradas ponderadas. Essa decisão terá efeito restrito ao próprio neurônio. Em modelos simples de redes neurais, a função de ativação pode ser a própria função soma das entradas ponderadas do neurônio. Em modelos mais complexos, a função de ativação possui um processamento atribuído. Esse processamento pode usar o valor prévio de saída como uma entrada para o próprio neurônio. Após ter sido processado pela função de ativação, é então passado pela função de transferência que produzirá o valor de saída do neurônio.

FALTA UMA LIGAÇÃO PARA EXPLICAR NEURÔNIO PARA APRENDIZADO ONLINE

# Perceptron de ε-Raio Fixo

Para construir o algoritmo, aplica-se as ideias da seção anterior à função de perda , restringindo a classe de funções C a funções lineares .  Assim, a condição para atualizar a hipótese após o exemplo é:

 (2.3.1.2)

Para a regra de correção, o gradiente na equação (2.3.1.1) é tomado em relação aos parâmetros (w, b) que compõem a função. Por isso:

(2.3.1.3)

em que, para . Esse algoritmo é chamado de Perceptron de ε- Raio Fixo (εPRF). Um algoritmo similar foi proposto porKivinen, A Smola e Williamson (2004), usando uma função de perda semelhante. Os algoritmos são equivalentes quando o parâmetro ν, usado porKivinen, A Smola e Williamson (2004), é definido como zero. O algoritmo εPRF é apresentado em detalhes no Algoritmo 1.

Algoritmo 1: εPRF em variáveis primais.

input :

output: (w, b)

1

2 repeat

3 for i = 1, . . . , m do

4 if then

5

6

7 t ← t + 1

  until nenhum erro foi cometido ou t > T

8 return

# Prova de convergência

A prova de convergência desenvolvida aqui segue os passos do teorema da convergência do algoritmo perceptron apresentado por Novikoff (1963). Ela garante que o εPRF convergirá em um número finito de iterações. Para o teorema a seguir define-se: R := maxi∈{1,...,m} ||xi ||, M := maxi∈{1,...,m} e m := mini∈{1,...,m} , em que.

Teorema 2.2.1.1. (Convergência εPRF): Dado um conjunto de treinamento e considerando uma solução (w\*,b\*), com um tubo de tamanho contendo os dados, o número de correções feitas pelo εPRF é limitada por

demonstração. Esta prova é construída de maneira similar ao teorema da convergência do perceptron. Seja o vetor normal ao hiperplano e o bias após a t-ésima correção. Suponha que essa correção ocorre para o i-ésimo exemplo. Lembre que a condição para um erro no i-ésimo exemplo é dado pela equação (2.2.1.2), que é equivalente a:

de onde tem-se:

(2.2.1.4)

Além disso, note que para uma solução ótima (w\*,b\*) e ε\* tem-se:

-

-

+ -

(2.2.1.5)

(2.2.1.6)

Usando o lado esquerdo de (2.2.1.5) e o lado direito de (2.2.1.6) leva à:

,

(2.2.1.7)

para qualquer que seja o valor de. A partir da correção dada pela equação (2.2.1.3) e usando a equação (2.2.1.4) tem-se o seguinte:

(2.2.1.8)

+

Ainda, a equação de correção dada por (2.2.1.2) leva a seguinte equação para o produto interno :

(2.2.1.9)

em que a expressão foi aplicada recursivamente e dado o fato que. Agora, combinando as equações (2.2.1.8) e (2.2.1.9) e aplicando a inequação de Cauchy-Schwarz tem-se:

de onde segue que

# Perceptron de ρ-raio fixo

Para construir o algoritmo de regressão ortogonal, considera-se a função de perda ρ- insensível apresentada na Seção 2. Seguindo uma derivação análoga, a condição para atualizar a hipótese após examinar o exemplo é:

A regra de correção correspondente tem a seguinte forma:

+

(2.2.1.10)

+

em que é dado por

(2.2.1.11)

:=

Esse algoritmo recebe o nome de Perceptron de ρ-Raio Fixo (ρPRF). Ele é apresentado em detalhes no Algoritmo 2.

Algoritmo 2: ρPRF em variáveis primais.

 input :

 output: (w, b)

1

2 repeat

3 for i = 1, . . . , m do

4 if then

5

6

7 t ← t + 1

  until nenhum erro foi cometido ou t > T

8 return

# SUPORTE VETOR MACHINE ONLINE

Revisando a Máquina Vetor de Suporte, o problema da modelagem de dados empírica é pertinente para muitas aplicações de engenharia. Na modelação de dados empíricos um processo de indução é utilizado para construir um modelo de sistema, a partir da qual espera-se deduzir respostas do sistema que ainda têm de ser observada. Em última análise, a quantidade e a qualidade das observações governar o desempenho deste modelo empírico. Por seus dados de natureza observacional obtido é finito e amostrado; normalmente esta amostragem é não uniforme e devido à natureza dimensional elevada do problema os dados vai formar apenas uma distribuição esparsa no espaço de entrada. Consequentemente o problema é quase sempre mal colocados (Poggio et al., 1985), no sentido de Hadamard (Hadamard, 1923). abordagens tradicionais redes neurais têm sofrido com dificuldades generalização, a produção de modelos que podem overfit os dados. Esta é uma consequência da algoritmos de otimização usado para a seleção de parâmetros e as medidas estatísticas usadas para selecionar o modelo de "melhor". As fundações de Support Vector Machines (SVM) têm desenvolvido por Vapnik (1995) e estão ganhando popularidade devido a muitos atraente características e desempenho empírico promissor. A formulação incorpora a Estrutural Risco Minimização princípio (MRE), que tem sido mostrado para ser superior, (Gunn et al., 1997), a princípio tradicional empírica Minimização de Risco (ERM), empregado por redes neurais convencionais. SRM minimiza um limite superior sobre o risco esperado, em oposição a MTC que minimiza o erro nos dados de treino. É esta diferença que equipa SVM com uma maior capacidade de generalização, que é o objetivo em estatística aprendizagem. SVM foram desenvolvidos para resolver o problema de classificação mas recentemente, eles foram estendidos para o domínio de problemas de regressão (Vapnik et al., 1997). Na literatura a terminologia para SVMs pode ser um pouco confuso. O termo SVM é tipicamente usado para descrever classificação com métodos de vetor de suporte e de vetores de suporte regressão é usado para descrever a regressão com os métodos de apoio de vetor.

Neste relatório o termo SVM irá referir-se a ambos os métodos de classificação e regressão, e os termos Suporte Vetor de Classificação (SVC) e Suporte Vetor de Regressão (SVR) será usado para a especificação.

Os problemas de séries temporais podem ser trazidos para regressão de vetores de suporte (SVR) problemas [2]. Em SVR, a idéia básica é mapear os dados em uma dimensão mais elevada espaço de características através de um mapeamento não linear e, em seguida, fazer de regressão linear neste espaço. Portanto, a regressão aproximada aborda o problema de estimar uma função com base em um determinado conjunto de dados (é o vetor de entrada, é o valor desejado). A Máquina Vetor de Suporte aproxima a função com a forma

Onde são os dados no espaço de recursos, e b é coeficientes. Eles podem ser estimados minimizando as funções de riscos regularizadas

(2)

Onde é a chamada função de perda de medindo os erros aproximados entre saída esperada e a saída calculada e C é uma constante regularização determinando a troca entre o erro de treinamento e o desempenho generalizado. O segundo termo, , é usado como uma medida de função planicidade. Introdução de variáveis lentas conduz (2) para a seguinte função restrita

Minimizar

Embora a função não linear é geralmente desconhecido  todos os cálculos relacionados com pode ser reduzida para a forma , que satisfaz as condições de Mercer [16,17]. Então, a equação (1) torna-se a forma explícita

(5)

Em (5), multiplicadores de Lagrangee satisfaz a igualdade. Esses vectores com são chamados vetores de suporte, as quais contribuem para a solução final.

O modelo SVM online apresentado cuja distinção primária do modelo convencional SVM centra-se sobre a forma de dados fornecidos, isto é, os dados são fornecidos em sequência para o modelo SVM on-line apresentado, enquanto eles são fornecidos em lotes para o modelo SVM convencional. No modelo proposto, sem perda de generalidade, é suposto que os dados de treinamento iniciais estabelecidos com amostras . A função kernel ótima com o parâmetro ótimo (por exemplo, o modelo ótimo é estabelecido baseado nos atuais dados ).

Para alguma base no espaço de Hilbert H, se , então . (cos(x,y)) refere-se à função cosseno do ângulo entre o vector x e y, eNa abordagem apresentada, uma sequência do vetor, é primeiramente construído, e então uma sequência de vetor ortogonal

(6)

com = classificação pode ser obtido por procedimenlo ortogonalização do Schmidt. Porque   (é o vetor normal do hiperplano ortogonal) é uma base ortogonal em e cada pertence a

(7)

consequentemente,

(8)

Porque todas as computações podem ser transformadas para o formato kernel, o parâmetro kernel ótimo pode ser obtido por minimizar . Quando uma observação nova chega (amostra ), é necessário determinar se trabalha bem com o modelo atual. Se por exemplo, , o modelo não for alterado, caso contrário o modelo ótimo é selecionado abaixo destas amostras . Para a conveniência das computações, é calculado na sequência. Aqui, denota

onde é algum índice em . Uma vez que o modelo ótimo sucessivo é obtido na base no modelo ótimo de prosseguir, é necessário apenas uma etapa de computação para determinar o kernel otimizado correspondente ao mínimo de um problema de otimização.

O custo de computação conjunto, em cada ciclo não irá aumentar em comparação com o treinamento Máquina Vetor de Suporte. Além disso, a coleta de dados não é contínua, isto é, existe sempre um intervalo entre medindo dois conjuntos de dados para aplicações reais.

Assim, o modelo apresentado online pode realizar seleção ótima do modelo e encontrar a melhor solução em simultâneo (Ou, digamos, pelos dados sequencial, ele seleciona o espaço do modelo ótimo no início, e, em seguida, encontra o ideal aproximação no espaço do modelo selecionado, de forma recursiva).

A ideia principal do modelo Máquina Vetor de Suporte online pode ser como a seguir:

onde

Passo 2

1. Para cada

Onde

(2) Então o Kernel ótimo é

Passo 3

(1) Se então calcula

# Suporte vetor de regressão

SVM também pode ser aplicado a problemas de regressão pela introdução de uma alternativa função de perda, (Smola, 1998). A função de perda deve ser modificado para incluir uma distância medida. Figura 5.1 ilustra quatro possíveis funções de perda.

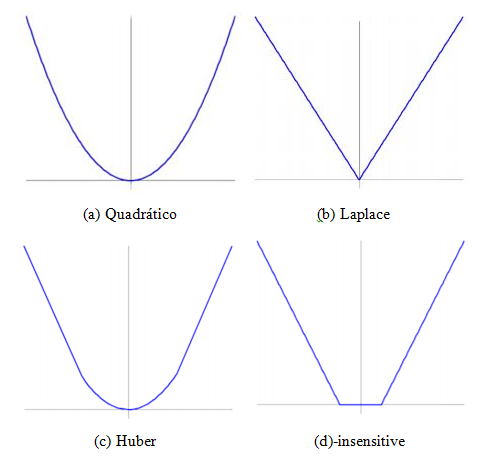


Figura 3: Função de perda.  
Fonte:Elaboração própria.

A função de perda na Figura 2.4 (a) corresponde ao erro de mínimos quadrados convencional critério. A função de perda na Figura 2.4 (b) é uma função de perda de Laplace que é menos sensível a outliers do que a função de perda quadrática. Huber proposta a função de perda na Figura 2.4 (c), como uma função de perda robusto que tem propriedades ótimas quando a base subjacente a distribuição de dados é desconhecido. Estas três funções de perda não produzirá sparseness nos vetores de suporte. Para abordar esta questão Vapnik proposta a função de perda de Figura 2.4 (d) como uma aproximação à função de perda de Huber que permite um conjunto esparsovectores de suporte para ser obtido.

1. **TECNOLOGIAS E TRABALHOS RELACIONADOS**

# TECNOLOGIAS EXISTENTES

O Waze (Suporte Google, 2016) é um aplicativo que compartilha informações em tempo real sobre o trânsito, acidentes, policiais, vias interditadas e condições climáticas. O aplicativo coleta as informações adicionadas pelo usuário e as analisa imediatamente para fornecer aos outros usuários (Wazers) a melhor rota para seus destinos, 24 horas por dia. Dirigir com o aplicativo aberto é uma maneira de contribuir com sua melhoria, quanto mais pessoas dirigirem com o Waze aberto, melhor será a navegação. As informações do percurso são utilizadas para calcular a velocidade média, encontrar erros e melhorar o traçado das vias, mapas e navegação além de aprender os sentidos de condução e conversões. O Waze funciona melhor nas suas rotas diárias e não há navegação disponível para o transporte público, bicicletas e caminhões.

Para que o aplicativo possa fornecer com precisão a rota mais rápida ou mais curta, de acordo com suas configurações, necessita de dados precisos de todos os segmentos e rotas vizinhas. O Waze coleta dados de cada trecho dirigido com o aplicativo aberto.  Na próxima vez em que uma via específica for percorrida, o aplicativo saberá como comparar os dados entre cada rota e saberá sugerir a rota ideal, conforme as preferências do usuário.

Algumas vezes, existem rotas melhores do que as sugeridas, problemas de roteamento como estes acontecem porque o sistema funciona em tempo real e usa a média dos dados estatísticos, por isso, o aplicativo sugere que algumas vezes, o usuário siga o caminho sugerido para que o Waze possa coletar novas informações e usá-las nos algoritmos de roteamento.

Quando o Waze não oferece a melhor rota, pode ser que houve um erro no mapa em algum lugar ao longo da rota preferida do usuário ou ausência de informações corretas de velocidade e trânsito na rota que está tentando enviar instante.

Para que todas as funções do Waze possam operar, é necessária uma conexão de dados ativa no dispositivo móvel para aprimorar com informações em tempo real sobre o trânsito e garantir um mapa atualizado. Se houver uma conexão intermitente, os dados serão obtidos dos servidores para alertas de trânsito e perigos, mas pode não ser capaz de fornecer informações confiáveis. Além disso, se o Waze não conseguir restabelecer a conexão com os servidores, o usuário não conseguirá enviar alertas, pois estes ou problemas não são armazenados no mapa para um posterior envio.

O Google Maps é um serviço de pesquisa e visualização de mapas e imagens de satélite da Terra gratuito na web fornecido e desenvolvido pela empresa Google.

Atualmente, o serviço disponibiliza mapas e rotas para qualquer ponto em vários países como nos Estados Unidos, Canadá, União Europeia, Austrália e Brasil. Disponibiliza imagens de satélite do mundo todo, com possibilidade de um zoom nas grandes cidades, como Nova Iorque, Paris, São Paulo, Rio de Janeiro e Brasília. Nos Estados Unidos o Google fez uma parceria com órgãos públicos, que incluirão as linhas de trem americanas e seus cruzamentos com rodovias ao Google Maps.

Juntamente com o lançamento da versão brasileira do Google Maps, a empresa introduziu o Local Business Center, ferramenta que permite com que qualquer empresa faça seu cadastro e seja encontrada no Google Maps. No cadastro as empresas podem preencher seus dados cadastrais, horário de atendimento, formas de pagamento, logotipo e fotos.

Com uma conta Google, é possível destacar as rotas, pontos e áreas, gerar comentários e compartilhar os respectivos links de acesso ao mapa criado. Também é possível gerar um arquivo KML para integração com o Google Earth.

O Google Maps sincroniza as pesquisas realizadas no computador para smartphones Android possibilitando o download do mapa, para uso “off-line”. O aplicativo recalcula o trajeto dinamicamente com base nos próximos padrões de trânsito, ajudando o usuário a evitar congestionamentos.

# TRABALHOS RELACIONADOS

QUAIS FERRAMENTAS DE BUSCA FORAM UTILIZADAS? Em [X] foi usado …

No entanto, os dois trabalhos mais relacionados ao apresentado nesta proposta são descritos a seguir.

Em (WIBISONO, A. et al., 2015) extração de informações através de sensores distribuídos tem sido amplamente utilizadas para obter informações conhecimento de várias regiões ou zonas. Veículo de extração de dados de tráfego é uma das maneiras de coletar informações a fim de

obter as informações sobre condição de tráfego. Essa pesquisa pretende prever e visualizar as

condições de tráfego em uma determinada região da estrada. Os dados de tráfego foram obtidos a partir do Departamento de Transportes do Reino Unido. Estes dados são recolhidos usando centenas de sensores durante 24h. Assim, o tamanho dos dados é muito grande. A fim de obter o comportamento da condição de tráfego, é necessário analisar o enorme conjunto de dados, que se obteve a partir dos sensores. As utilizações dos métodos de extração convencionais de dados não são suficientes para a utilização, devido ao processo de conhecimento construção de ponta que deve armazenar dados temporários na memória. O fato de que os dados vão se tornando maior ao longo do tempo, é preciso encontrar um método que poderia se adaptar automaticamente para processar dados sob a forma de fluxos. O método chamado FIMT-DD (Fast Incremental Modelo Trees-Deriva Detection) foi usado para analisar e prever o grande conjunto de dados de tráfego. Baseado no sistema de previsão de que foi desenvolvido, foi também visualizada a previsão da condição de fluxo de tráfego dentro do ponto de sensor gerado no mapa real de simulação. DIFERENÇA??

No artigo (WENJIAN, W. et al., 2007), desenvolve um modelo SVM online para prever níveis de poluentes do ar em uma série temporal avançada com base no banco de dados poluente do ar monitorada no centro da cidade de Hong Kong. A comparação experimental entre o modelo SVM on-line e o modelo SVM convencional (modelo SVM não online) demonstra a eficácia e eficiência na previsão de parâmetros de qualidade do ar com diferentes séries temporais. Nos problemas de previsão de séries temporais, tem havido vários modelos de previsão dos dados, mas o desenvolvimento de um modelo mais preciso é muito difícil por causa de elevadas relações não lineares e não estáveis ​​entre os dados de entrada e de saída. Quase todos os modelos em mãos não são aplicáveis online, embora a previsão online, especialmente para os parâmetros de qualidade do ar previsão, tem um significado muito importante para reais aplicações do mundo. Uma máquina de vetor de suporte (SVM), como uma ferramenta nova e poderosa de aprendizagem de máquina, pode ser usada para séries temporais predição e tem sido relatado para um bom desempenho de alguns resultados promissores. Em (LEE, J. et al., 2015) são descritos novos conhecimentos entre congestionamento do tráfego e tempo usando tecnologia de processamento de um grande conjunto de dados armazenados. Os fatores climáticos como temperatura e precipitação em zonas residenciais e destinos turísticos afetam o fluxo de tráfego nas estradas circundantes. As mudanças no congestionamento do tráfego devido ao clima são avaliadas usando linear múltipla. A análise de regressão cria um modelo de predição e previsão de congestionamento do tráfego em uma base diária. Para a análise de regressão, foram usados 48 fatores de previsão meteorológica e seis variáveis ​​dummy(variáveis utilizadas para representar numericamente uma variável qualitativa) para expressar os dias da semana. A regressão linear múltipla do modelo final é então proposta com base nas três etapas de análise de (I) a criação de um modelo de regressão completa, (II) a remoção das variáveis, e (III) análise dos resíduos. Para verificar a sua previsibilidade, o modelo avalia o congestionamento do tráfego, prever o congestionamento do tráfego com o congestionamento de tráfego real. Usando o método de avaliação erro médio percentual absoluto, mostra que o modelo de regressão linear múltipla final tem uma precisão da previsão de 84,8%. DIFERENÇA??

1. **METODOLOGIA DA PESQUISA**

A metodologia deste trabalho consiste na continuidade de estudos de referencial bibliográfico que aborde a aplicação dos algoritmos de aprendizado online para problemas de regressão para predição de horário de saída de rotas. Serão abordados os algoritmos Regressão Linear Online, RNA (perceptron) online e Maquina Vetor de Suporte online.

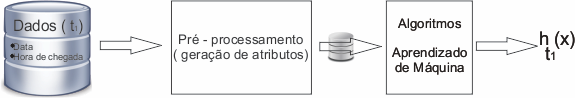
Na Figura 4 é apresentado um diagrama esquemático das atividades a serem executadas para atingir o objetivo dessa pesquisa. 

Figura 4: Diagrama esquemático das atividades.  
Fonte:Elaboração própria.

**1 - Coleta dos dados:** Nessa etapa, os dados são coletados da Tomtom. Na Tabela 1 é mostrada uma amostra de dados que foram coletados para preparação dessa metodologia.

por exemplo, para calcular o tempo estimado de uma viagem. O tempo estimado (variável dependente) é o resultado de muitas variáveis independentes: a data de saída, a hora de chegada, o local de origem e o local de destino, o modelo é criado com base no dados da tabela 2.1 previamente capturados.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Distância(m)** | **Tempo Total(s)** | **Tempo de Atraso(s)** | **Data e hora de partida** | **Data e hora de chegada** |
| 5946 | 1098 | 0 | 2015-12-23T20:37:49-02:00 | 2015-12-23T20:56:06-02:00 |
| 5946 | 1078 | 0 | 2015-12-23T20:52:56-02:00 | 2015-12-23T21:10:53-02:00 |
| 5946 | 1052 | 0 | 2015-12-23T21:08:03-02:00 | 2015-12-23T21:25:35-02:00 |
| 5946 | 1026 | 0 | 2015-12-23T21:23:09-02:00 | 2015-12-23T21:40:15-02:00 |
| 5946 | 983 | 0 | 2015-12-24T21:47:23-02:00 | 2015-12-24T22:03:45-02:00 |
| 5946 | 978 | 0 | 2015-12-24T22:02:27-02:00 | 2015-12-24T22:18:44-02:00 |
| 5946 | 978 | 0 | 2015-12-24T22:17:31-02:00 | 2015-12-24T22:33:48-02:00 |
| 5946 | 978 | 0 | 2015-12-24T22:32:35-02:00 | 2015-12-24T22:48:52-02:00 |
| 5946 | 978 | 0 | 2015-12-24T22:47:39-02:00 | 2015-12-24T23:03:56-02:00 |
| 5946 | 978 | 0 | 2015-12-24T23:02:43-02:00 | 2015-12-24T23:19:00-02:00 |
| 5946 | 978 | 0 | 2015-12-24T23:17:46-02:00 | 2015-12-24T23:34:03-02:00 |
| 5946 | 978 | 0 | 2015-12-24T23:32:51-02:00 | 2015-12-24T23:49:08-02:00 |
| 5946 | 978 | 0 | 2015-12-24T23:47:54-02:00 | 2015-12-25T00:04:11-02:00 |
| 5946 | 978 | 0 | 2015-12-25T00:02:57-02:00 | 2015-12-25T00:19:14-02:00 |
| 5946 | 978 | 0 | 2015-12-25T00:18:01-02:00 | 2015-12-25T00:34:18-02:00 |
| 5946 | 978 | 0 | 2015-12-25T00:33:04-02:00 | 2015-12-25T00:49:21-02:00 |
| 5946 | 978 | 0 | 2015-12-25T00:48:08-02:00 | 2015-12-25T01:04:25-02:00 |
| 5946 | 978 | 0 | 2015-12-25T01:03:12-02:00 | 2015-12-25T01:19:29-02:00 |
| 5946 | 978 | 0 | 2015-12-25T01:18:15-02:00 | 2015-12-25T01:34:32-02:00 |
| 5946 | 978 | 0 | 2015-12-25T01:33:19-02:00 | 2015-12-25T01:49:36-02:00 |
| 5946 | 978 | 0 | 2015-12-25T01:48:22-02:00 | 2015-12-25T02:04:39-02:00 |

Tabela 1: Uma amostra dos dados disponíveis para o trabalho.   
Fonte: Dados coletados da Tomtom.

2 - Pré-processamento dos dados: A etapa de pré-processamento, no processo de descoberta de conhecimento – KDD (Knowledge Discovery in Databases), compreende a aplicação de várias técnicas para captação, organização, tratamento e a preparação dos dados. É uma etapa que possui fundamental relevância no processo de KDD. Compreende desde a correção de dados errados até o ajuste da formatação dos dados para os algoritmos de mineração de dados que serão utilizados. Os atributos se é feriado ou não, véspera de feriado ou não, final de semana serão acrescentados como atributo 0 (não) ou 1(sim).

Os atributos Distância, Tempo Total, Tempo de Atraso, Data e Hora de partida, Data e Hora de chegada, Rota, Terça/Quarta/Quinta, Segunda, Sexta, Sábado, Domingo, Pós-feriado, Pré-feriado

ESCREVER OS ATRIBUTOS A SEREM EXTRAÍDOS DO CONJUNTO DE DADOS E COMO O PROBLEMA SERÁ MODELADO – EX: CONSTRUIR CLASSIFICADOR PARA CADA ROTA.

3 - Indução dos Modelos:

4 - Avaliação dos Modelos:

1. **ATIVIDADES E CRONOGRAMA**

As seguintes atividades serão realizadas:

1 - Levantamento Bibliográfico

2 – Coleta e pré-processamento de dados

3 – Indução e Avaliação dos Modelos

4 - Escrita e Submissão do Artigo

5 - Escrita e Dissertação

7 - Defesa da Dissertação

8 - Conclusão

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2015 | | | | | | | | | | | | 2016 | | | | | |
|  | Meses | | | | | | | | | | | | | | | | | |
| Atividades | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. **considerações finais**

A presente proposta teve como objetivo analisar as tecnologias existentes, apresentar os algoritmos que serão utilizados na execução para fazer a comparação e identificar o que terá o melhor resultado.

# REFERÊNCIAS

ALMEIDA, Maria Aparecida Fernandes. **Introdução ao Estudo das Redes Neurais Artificiais.**2000. Disponível em: <https://intranet.dcc.ufba.br/pastas/mecateam/material\_de\_estudo/ia/iaconexionista/rna.pdf>. Acesso em: 23 maio 2016.

BROWNLEE, Jason. **A Tour of Machine Learning Algorithms.**2013. Disponível em: <http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>. Acesso em: 02 maio 2016.

DIA A DIA EDUCAÇÃO (Brasil). Secretaria da Educação. **Esquema de um neurônio.**2016. Disponível em: <http://www.biologia.seed.pr.gov.br/modules/galeria/detalhe.php?foto=256&evento=3>. Acesso em: 22 maio 2016.

FIAT, A. **Online Algorithms:**The State of the Art (Lecture Notes in Computer Science). Springer, 1998.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social.**São Paulo: Atlas, 1999.

KIVINEN, J; A SMOLA; WILLIAMSON, R. **Online learning with kernels.**Ieee Transactions On Signal Processing, 2004. 52 v.

NOVIKOFF, A B. **On convergence proofs for perceptrons.**In: Proceedings of the Symposium on the Mathematical Theory of Automata, 1963. 12 v.

SMOLA, A J; SCHÖLKOPF, B. **On a kernel–based method for pattern recognition, regression, approximation and operator inversion.**Algorithmica, 1998.

TAFNER, Malcon Anderson. **O Que São as Redes Neurais Artificiais.**1998. Disponível em: <http://www.cerebromente.org.br/n05/tecnologia/rna.htm>. Acesso em: 26 maio 2016.

VAPNIK, V. **The Nature of Statistical Learning Theory.**New York: Springer, 1995.

VAPNIK, V. **Statistical Learning Theory.**New York: Wiley, 1998.

WIBISONO, Ari et al. **Traﬃc big data prediction and visualization using Fast IncrementalModel Trees-Drift Detection (FIMT-DD).**2015. Disponível em: <https://www.researchgate.net/publication/284121757\_Traffic\_Big\_Data\_Prediction\_and\_Visualization\_using\_Fast\_Incremental\_Model\_Trees-Drift\_Detection\_FIMT-DD>. Acesso em: 26 maio 2016.

WIKIPÉDIA. **Google Maps.**2016. Disponível em: <https://pt.wikipedia.org/wiki/Google\_Maps>. Acesso em: 03 jun. 2016.