

ANÁLISE DE DEMANDA VIA INTELIGÊNCIA ARTIFICIAL NO RESTAURANTE UNIVERSITÁRIO DO INSTITUTO DE CIÊNCIA E TECNOLOGIA DA UNIFESP

1st Douglas Diniz Landim

Graduando em Ciência da Computação (disciplina de Inteligência Artificial)

Universidade Federal de São Paulo (Instituto de Ciência e Tecnologia)

São José dos Campos, Brasil

ddlandim@unifesp.br

Resumo—O presente trabalho tem como objetivo a comparação com o método de análise da previsão de vendas do restaurante universitário da Unifesp, previamente feita pelo autor deste projeto com aplicação de métodos estatísticos, e atualmente com métodos de aprendizado de máquina. Tal análise por aprendizado de máquina foi apontado como relevante solução no fim da análise estatística do trabalho anterior. A temperatura da região onde se localiza o campus do restaurante foi analisada por recorrência via bootstrap como um fator que exerce impacto sobre as vendas do restaurante em certos períodos do semestre, e neste trabalho de conclusão de curso serão obtidas novas variáveis e um intervalo maior de amostragem na análise da previsão de demanda de refeições do restaurante em um novo modelo com aprendizado de máquina a fim de que seja obtido um modelo de previsão viável para evitar super-projeção de demanda com consequência de desperdício de alimentos, ou subprojeção com consequência de docentes ou discentes sem refeições.

Index Terms—Redes Neurais Artificiais. Previsão de demanda. Aprendizado de Máquina. Inteligência Artificial. Perceptron Múltiplas camadas.

I. INTRODUCTION

A previsão da demanda é o fator principal da eficiência de qualquer modelo de processamento do tipo entrada-saída, onde a sua saída deve atender uma demanda não determinística. É necessário prever a demanda para projetar e aperfeiçoar o processamento e a entrada deste modelo. Portanto, a previsão de demanda é um ponto de extrema importância para qualquer empresa e, sendo adequada, permite o ajuste de todo o mecanismo de operações da empresa para atendê-la com a melhor eficiência possível, maximizando lucros, minimizando perdas e, principalmente, atendendo todas as necessidades do cliente.

Todo restaurante universitário enfrenta problemas de previsão de demanda de refeições e prejuízos com a falta de vendas e ou o descarte de refeições não vendidas. Um dos grandes problemas enfrentados hoje no mundo é a elevação dos preços dos alimentos. O valor, além de monetário, é moral pois a alimentação é o recurso primitivo de base da humanidade que hoje ainda enfrenta um difícil acesso a este em muitas regiões carentes. O descarte de alimentos, provocado por suas

limitações e durabilidade, gera não somente prejuízos monetários, mas ambientais e morais. Isso tem causado preocupações para a população em geral e também para empresas como restaurantes que sofrem diretamente os reflexos da variação no preço dos alimentos e na demanda. Atualmente, o restaurante universitário do ICT - UNIFESP não possui um sistema que ajude na gestão de compras dos alimentos e sofre dos problemas acima relatados.

No restaurante universitário do ICT - UNIFESP, as refeições são fornecidas de segunda a sexta-feira. O caso particular de restaurantes universitários envolve um fluxo de demanda influenciado pelo dia da semana, visto que a demanda é influenciada pela quantia de alunos presentes na universidade, que por sua vez é influenciada pela grade de aulas determinada semestralmente por dia da semana. O caso de análise para este projeto foi motivado após informações de relevantes desperdício de alimentos.

Devido às condições burocráticas no ambiente do restaurante que compreendem fidelidade de contrato e exclusividade, pois o restaurante se encontra em localização que o faz ser o único provedor de refeições ao público do campus, a escolha dos parâmetros não será influenciada por muitos fatores externos como concorrência, acessibilidade do ponto, entre outros.

Outro ponto importante é a obtenção dos valores de venda; não foram escolhidas as vendas diretas do ponto de venda de tickets de refeição, e sim os dados de coleta da entrada do restaurante, que demonstram a real movimentação de público no restaurante em determinado dia.

O estudo da relação de vendas, temperatura, outras variáveis climáticas e do ambiente, já é comum em outros cenários; dentre eles, o de maior destaque é na demanda de energia elétrica. Os cenários de vendas de alimentos perecíveis ganha também destaque apesar de se encontrar investimentos maiores de previsão de demanda na indústria de produção de energia elétrica. O objetivo, tanto no cenário deste trabalho - o restaurante universitário, como em outros cenários é o mesmo: atender toda a demanda de consumo e evitar transtorno a qualquer consumidor pela falta do produto desejado e evitar

prejuízos de produção não consumida. Tais prejuízos impactam não só o fornecedor, mas o consumidor, visto que um fornecimento de produto e serviço com um bom planejamento de demanda poupa recursos ao produtor. Tais recursos podem ser investidos em melhor qualidade de produto e menor preço ao consumidor caso seja obtido um modelo de previsão viável para evitar sobrestimação de demanda com consequência de desperdício de alimentos, ou subestimação com consequência de docentes ou discentes sem refeições.

Tal problema tem sido impactante e frequente para o restaurante que informa que em alguns dias no mês passa por sobrestimação e desperdício superior a 30% do total produzido.

O modelo a ser analisado neste trabalho é o comportamento dos consumidores de um restaurante universitário (R.U), o qual precisa projetar sua compra de insumos e alocação de recursos na entrada de seu modelo de negócio, e projetar sua saída, que é a produção de refeições em quantidade numérica e inteira distribuída em função do tempo em dias. O consumo feito por alunos não se comporta de maneira determinística, já que este consumo é facultativo. Portanto, nem todos os alunos que estão na universidade se alimentarão no R.U. todos os dias.

De acordo com o contrato presente entre o R.U e a universidade, o mesmo deve atender totalmente à demanda do público, sendo multado se caso algum consumidor fique sem alimentação. Porém, este mesmo contrato não trata refeições que não são consumidas; logo, o R.U. deve lidar integralmente o prejuízo de refeições produzidas acima da demanda de consumo. O que o R.U. faz com essas sobras está fora do escopo deste trabalho.

Tais refeições fornecidas aos alunos são, em parte, subsidiadas pela universidade. O período de dados obtido para este trabalho compreende até Agosto de 2018 e segue um modelo de contrato antigo no qual o restaurante recebe R\$2,50 do aluno e R\$9,14 da universidade totalizando R\$11,64 por refeição.

De agosto em diante, iniciando no segundo semestre de 2018, a universidade subsidia R\$5,44 da refeição do aluno e o mesmo R\$2,50, o R.U. recebe o total de R\$7,94. Logo, esta previsão de demanda corresponde também aos interesses da administração do campus local, que periodicamente deve realizar uma alocação de recursos financeiros para subsidiar todas estas refeições consumidas.

II. DEFINIÇÃO DO PROBLEMA

O problema a definir neste trabalho é encontrar um modelo de previsão de demanda por meio de algoritmos de regressão via aprendizado de máquina.

III. JUSTIFICATIVAS

As atuais abordagens de previsão do restaurante universitário, que envolvem análise exploratória semanal e dedução subjetiva pelos responsáveis pelo restaurante são falhas por não serem calculadas considerando-se uma visão ampla de todo um histórico de dados de grande amostragem. Além disso, não cruzam outras informações como dados climáticos,

dados de calendário anual, feriados próximos, dentre outros. Essas falhas causam desperdício de alimentos por superestimação de demanda e prejuízo de refeições descartadas, e falta de atendimento por sub-estimação de demanda de consumo.

IV. OBJETIVOS

Construir um modelo utilizando uma Rede Neural Artificial para a previsão da demanda de refeições do restaurante universitário do ICT-UNIFESP

V. METODOLOGIA

A. Composição de dados e Pre Processamento

Por se tratar de um trabalho de previsão de demanda, este trabalho irá realizar a coleta e tratamento dos dados de consumo do restaurante universitário da Unifesp, a coleta de dados climáticos que influenciam em tal demanda como temperatura máxima do dia e precipitação em %, calcular os dados derivados das datas de coleta, sendo distância posterior e anterior em dias entre cada data de observação pois distâncias grandes entre observações sinalizam feriados prolongados ou recessos que podem impactar na frequência dos alunos em um último ou primeiro dia entre o recesso, estação do ano, valor categórico do dia da semana, e dados dos últimos 5 dias de venda para prever o dia seguinte. Tais estruturação de dados e análises exploratórias são fundamentadas por pós (JUNIOR, 2007), Capítulo 2.4, para previsões de demanda em geral.

O modelo de dados de venda estruturados recebe o acréscimo de dados de variáveis climáticas, como possíveis fatores de influencia no consumo, conforme ocorre no trabalho de previsão de demanda de energia elétrica fundamentado por (ALMEIDA, 2013; RUAS, 2012; SILVA, 2010). Os trabalhos de (JUNIOR, 2007) e (SILVA, 2010) fundamentam também uma classificação de análises dos dados estruturados indicando os métodos de previsão de demanda para este modelo quantitativo que são a Regressão Linear Múltipla e Redes Neurais Artificiais com Múltiplas Camadas de Perceptrons.

B. Estudo de Aprendizado de máquina

As análises de inteligência artificial, fundamentadas por trabalhos de previsão de demanda em R.U. da Universidade Federal de Viçosa, de acordo (LOPES, 2008) e na Universidade Estadual Paulista Júlio de Mesquita por (ROCHA JOSÉ CELSO; MATOS, 2011), concluem a segunda aplicação da técnica de redes neurais artificiais com o método de Perceptron de Múltiplas Camadas (PMC) fundamentado em (HAYKIN, 1994), incrementando-se na técnica a inclusão de variáveis climáticas utilizadas nos trabalhos de previsão de (ALMEIDA, 2013; RUAS, 2012; SILVA, 2010), modificando a topologia como em (LOPES, 2008) para receber tais dados.

C. Análise dos resultados

De acordo com (MARIANO, 2014), a comparação e a discussão dos resultados deve ser feita com base em medidas de avaliação de erro absoluto médio (EAM), erro quadrado médio (EQM) e raiz do erro quadrado médio (REQM).

VI. FUNDAMENTAÇÃO TEÓRICA

A. Análise Exploratória dos Dados

(JUNIOR, 2007) Cita que dados coletados em modelos de previsão possuem informações que quando são projetadas graficamente evidenciam comportamentos que em alguns casos podem ser visualizados e generalizados de forma subjetiva pelos gestores dos dados. Em todos os casos, a análise exploratória é necessária para selecionar o melhor método de análise que se enquadra neste comportamento.

Somente a análise exploratória não é o suficiente e realizada nos intervalos ou critérios incorretos pode comprometer seriamente as conclusões do comportamento dos dados, e que por sua vez pode comprometer seriamente a decisão dos gestores responsáveis por estes dados, no cenário de uma previsão de demanda. Isto ocorre atualmente no cenário de previsão da demanda de refeições do ICT-UNIFESP, onde a universidade e o estabelecimento que fornece as refeições não tem nenhum modelo de previsão de demanda.

De acordo com o gestor da atual empresa que fornece refeições no ICT, a análise utilizada para se prever as refeições é observar dentro da semana o dia anterior de consumo. Em variações de 300 para 450 refeições aproximadamente, isso tem provocado um desperdício médio de 150 refeições diárias. Em geral, de acordo com o restaurante, todos os dias o mesmo trabalha com um erro e um descarte de 30% das refeições que são trazidas e consumidas ao campus. Estima-se então que no período de 2011 - 01/08/2018 os estabelecimentos tenham tido um prejuízo de R\$1.885.938,40, e de 30% de R\$78.386,85 no atual período de 01/08/2018 - 31/10/2018 totalizando o montante R\$1.964.325,25. Aproximadamente 2 milhões de reais em prejuízo acumulado desde 2011.

B. Métodos de Previsão

(JUNIOR, 2007) Realiza uma revisão bibliográfica extensa abordando principais métodos de previsão de consumo sazonais, no cenário de uma indústria cosmética. Tais métodos estatísticos de previsão se dividem em 2 ramificações, sendo quantitativos ou qualitativos. Métodos qualitativos fazem um julgamento dos dados expostos sem um sistema de processamento analítico para se produzir novos modelos ou dados, eles são úteis para sistemas de agrupamento, clusterização ou classificação de dados, sem fornecer novas informações numéricas ou modelos preditivos.

Métodos quantitativos que é o foco deste trabalho, são analíticos e se baseiam em um modelo matemático para realizar previsões.

Para realizar tais previsões os métodos quantitativos necessitam de um histórico de dados, para analisar padrões em seu comportamento e prever o futuro que irá agir dentro deste padrão.

Estes métodos se ramificam em 2 tipos, as séries temporais e os modelos causais.

C. Métodos de previsão de Demanda

O autor supracitado também referencia métodos estatísticos especialmente selecionados para uma previsão de demanda,

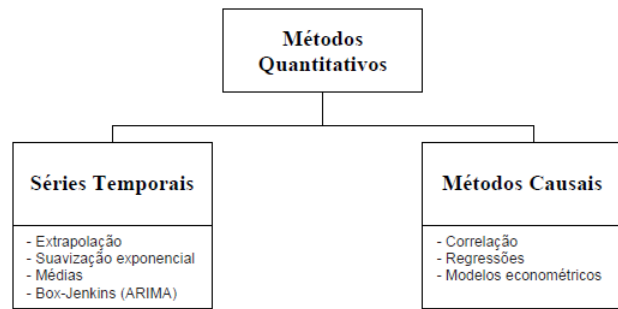


Figura 1. Tipos de métodos quantitativos Retirado de (JUNIOR, 2007).

com a atenção de que alguns métodos qualitativos foram criteriosamente selecionados para prever uma demanda industrial, onde geralmente são previstas pelos métodos quantitativos.

O comportamento dos dados deste trabalho, apesar de ter uma distribuição de datas em função do tempo se classificando em um modelo de série temporal, assume-se a hipótese que tem tal comportamento impactado por relações causais com outras variáveis como recesso acadêmico, feriados, eventos, precipitações intensas que causam trânsito local e impactam na logística e frequência do público, entre outras variáveis de causas menos aparentes.

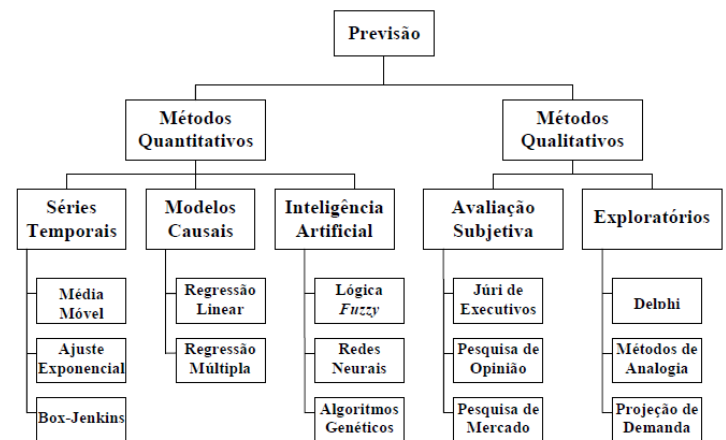


Figura 2. Métodos de previsão de demanda (JUNIOR, 2007).

D. Inteligência Artificial

A *Inteligência artificial*: "A inteligência artificial é o ramo da ciência da computação que se ocupa do comportamento inteligente."(LUGER, 2004). O termo "Inteligência Artificial" surgiu em meados de 1956 em uma conferência, chamada Dartmouth Summer Research Project on Artificial Intelligence, sediada nos Estados Unidos em Dartmouth College, Hanover, New Hampshire, a conferência foi organizada por John McCarthy e teve como proposta reunir matemáticos, cientistas da computação e pesquisadores buscando meios de como fazer com que as máquinas usem linguagem, abstrações de formulários e conceitos, para resolver problemas que eram

reservados aos humanos, e que consigam melhorar o seu próprio desempenho.

Sistemas de inteligência artificial buscam então resolver funções e problemas que seres humanos conseguem resolver melhor do que máquinas convencionais, usando sua capacidade de abstração e aprendizagem com o erro.

E. Neurônio Artificial

História e inspiração da Inteligência Artificial: Redes Neurais Artificiais são elementos de inteligência artificial da ciência da computação, inspirados no funcionamento do cérebro humano. São formadas por neurônios artificiais interconectados que são capazes de processar múltiplos valores de entradas, e reagir produzindo uma resposta relacionada à essas entradas. Como qualquer outro método de aprendizado de máquina, este modelo busca obter um aprendizado a partir dos dados de entrada recebidos, criando uma capacidade de generalização de problemas e assim buscam o objetivo principal de resolver novos problemas com este aprendizado.

Uma rede neural pode não produzir uma resposta esperada, resolvendo erroneamente um problema, assim como o cérebro humano tem limitações de aprendizado, que levam o homem a cometer falhas de decisões ou ações por um aprendizado mal treinado. Por isso as características fundamentais que tornam uma rede neural artificial em uma boa solução, é um bom planejamento de sua topologia, e método de treinamento, que serão explicados a seguir.

(BRYAN KOLB; ARIF, 2014) Demonstra em sua pesquisa que a propriedade fundamental do cérebro é a sua capacidade de mudar com uma grande variedade de experiências, inclusive lesões que provocam perdas de neurônios, abordando princípios de plasticidade cerebral, mostra também que a capacidade de aprendizado de mamíferos que sofreram uma redução de neurônios causada por lesões, pode ser restaurada não somente pela recuperação desses neurônios perdidos, mas sim por novas conexões e sinapses entre outros neurônios. Ou seja, a rede sofre uma mudança de topologia.

(ZORZETTO, 2012) Mostra em um experimento recente que o cérebro humano possui atualmente 86 bilhões de neurônios interconectados, e que sua capacidade de aprendizado e habilidades evolutivas também vêm aumentando em conjunto com seu número de neurônios e topologia de rede neural, que há 30 milhões de anos atrás tinha apenas 2,5 bilhões de neurônios e era apenas um animal arborícola quadrúpede.

Neurônio Biológico: O modelo de neurônio artificial surge então, com a busca da inteligência artificial de reproduzir o comportamento de aprendizado humano, reproduzindo computacionalmente seu elemento biológico principal de aprendizagem: O neurônio biológico. Um neurônio biológico (Figura 03) processa informações recebidas por meio de seus dendritos e processa-as em seu corpo celular, tal reação à esses estímulos recebidos gera um sinal de saída como resposta aos estímulos, enviado através do axônio. Esse estímulo é repassado como sinal de entrada, através de outros neurônios por meio de seus dendritos, e o ciclo se repete em uma vasta rede. Controlando essas conexões, pontos de contato entre a resposta de um

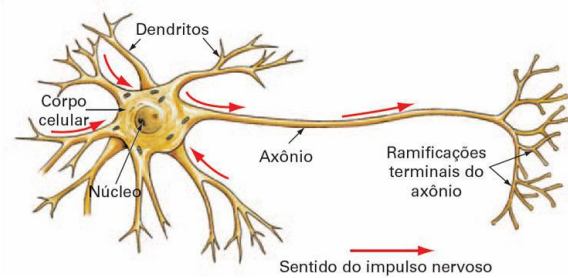


Figura 3. Neurônio Biológico, retirado de <https://pt.khanacademy.org/science/biology/human-biology/neuron-nervous-system/v/anatomy-of-a-neuron>.

neurônio e a entrada de outro, existem as sinapses. Elas funcionam como agentes que permitem a interação acontecer ou inibem, e são acionadas por um conjunto somatório de estímulos. Se tal somatório de estímulos for satisfatório, elas permite a transmissão de sinal elétrico pelo axônio até o dendrito de um neurônio vizinho, formando um ciclo de aprendizado em uma rede neural biológica.

Neurônio Artificial: Warren McCulloch e Walter Pitts em 1942 observando o neurônio biológico iniciam a busca de um modelo computacional do mesmo. McCulloch era psiquiatra e neuroanatomista e passou cerca de 20 anos refletindo e estudando sobre a representação do sistema nervoso, em 1942 ele convidou Pitts, que era matemático, para fazer parte das suas pesquisas. Em 1943 lançaram o artigo "A Logical Calculus of the Ideas Immanent in Nervous Activity," chegando em um modelo matemático do neurônio artificial:

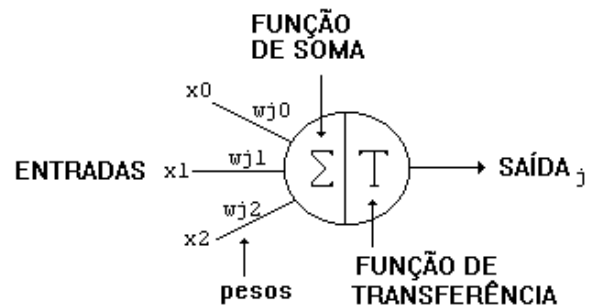


Figura 4. Neurônio Artificial, Retirado de <http://redesneuraisartificiais.blogspot.com/2010/10/o-primeiro-modelo-de-um-neuronio-criado.html>

Assim como no neurônio biológico, este modelo reage à um vetor de entradas X_0, X_1, \dots, X_n onde tem suas sinapses representadas por pesos numéricos, existe uma soma ponderada dessas entradas que é controlada por uma função de transferência ou função de ativação, determinando se essa soma é maior que um valor numérico. Se essa soma for satisfatória o neurônio é ativado emitindo um valor de saída 1, caso contrário se emite um valor de saída 0. Todo o funcionamento deste modelo então é reduzido a responder se

a soma recebida é maior que um valor numérico esperado.

Neste modelo simplista o neurônio consegue realizar operações lógicas. Na figura 05 os pesos w_1 e w_2 e o limiar T estão ajustados para responder à operação AND que resumidamente à uma operação lógica que somente produz um valor positivo 1 na resposta de saída se ambos os valores de entrada forem 1.

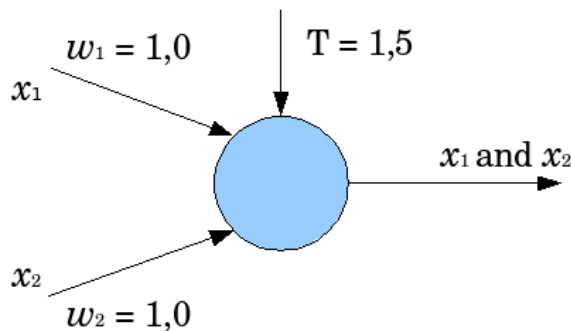


Figura 5. Neurônio Artificial - Operação AND, Retirado de <http://redesneuraisartificiais.blogspot.com/2010/10/o-primeiro-modelo-de-um-neuronio-criado.html>

O neurônio irá comparar se a soma das entradas X multiplicadas pelos seus devidos pesos W , será menor que o limiar T . Ele somente irá produzir uma resposta positiva (1) se essa soma for maior. É fácil verificar então na imagem, que o resultado dessa soma só será positivo, se ambas as entradas forem 1, onde a soma será $2 > 1,5$.

Tabela I
OPERAÇÃO LÓGICA AND

X_1	X_2	Saída
0	0	0
0	1	0
1	0	0
1	1	1

O perceptron (MARIANO, 2014) possui apenas uma camada de entrada e saída, a saída utiliza como função de ativação a função degrau u que define quando o neurônio emitirá o sinal lógico 1 ou quando emitirá o sinal lógico 0.

O sinal de saída do perceptron entende-se então por: $y = \delta(\sum_{i=1}^n X_i W_i + b)$

- X_i - sinais de entrada do neurônio;
- W_i - pesos sinápticos do neurônio;
- b - bias ou limiar de ativação;
- $\delta(\cdot)$ - função de ativação;
- y - sinal de saída do neurônio.

Em algumas literaturas o bias pode ser reduzido a um peso W_0 com entrada X_0 fixo em 1 no neurônio, a representação gráfica da topologia pode mudar, mas no somatório de saída, o cálculo continua o mesmo.

A função de ativação δ se apresenta de forma linear ou não linear, determinando a saída de um neurônio a partir do

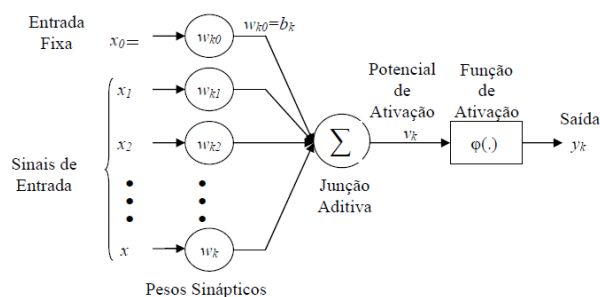
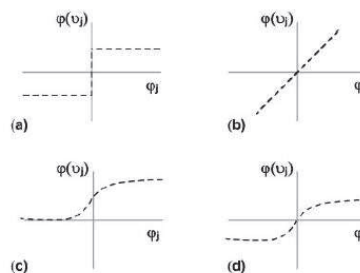


Figura 6. Neurônio Artificial Perceptron, retirado de (JUNIOR, 2007) ()

seu potencial de ativação. As funções de ativação podem ser observadas na figura 09, sendo a de maior popularidade de utilização, a sigmoide.



Tipos de função de ativação: (a) função degrau, (b) função linear, (c) função sigmoide, (d) função tangente hiperbólica

Figura 7. Funções de Ativação retirado de (MARIANO, 2014)

As funções de ativação dá a capacidade do perceptron quando conectado em rede (explicado no próximo capítulo) de resolver problemas lineares e não lineares, agregando adaptação e improviso ao resolver programas que não estão contidos em seus dados de alimentação.

Perceptron - Treino Supervisionado: (ALMEIDA, 2013) cita que o processo de aprendizado do perceptron pode ocorrer de forma supervisionada quando o neurônio deve aprender a relacionar um conjunto observado de variáveis à um valor observado de saída deste mesmo conjunto. O neurônio recebe os sinais de entrada X_i e produz uma saída Y_i através do combinador linear e função δ de ativação; compara essa saída Y_i com a observação i obtida do conjunto de dados (ponto chave do processo de treino supervisionado) e por fim essa comparação irá gerar um erro e .

* Critério de parada De acordo com algum critério adotado em cada contexto de aplicação do perceptron, este erro pode ser aceitado e o neurônio mantém o valor de seus pesos W_i que impactam na saída Y_i desejada. A aprendizagem do neurônio também pode atingir um critério de parada após N épocas de treinamento, diversos critérios de paradas podem ser combinados.

* Época de treinamento Se o critério de parada não for aprovado, uma nova época de treinamento (ou repetição de interação) é iniciada mesmos valores X_i e Y_i passados, os valores de pesos W_i são reajustados com a taxa de aprendizagem, buscando o objetivo de se obter um erro menor.

* Taxa de aprendizagem Esse reajuste de pesos W_i denomina-se taxa de aprendizagem, α que pode ter valores de escolha livre ao contexto de aplicação do perceptron, para reajustar estes pesos W_i . Assim que a taxa de aprendizagem ajusta os pesos, uma nova época $N+1$ de treinamento está se iniciando buscando novamente um erro menor. Ressaltando então que o critério de parada pode ser acionado e interromper o reinício do processo, se for estipulado um limite para o valor de N combinado ou não com um limite para o erro.

F. Funcionamento e treino do neurônio perceptron

Em (ALMEIDA, 2013) verifica-se de forma procedural, o processo do treino supervisionado do perceptron, com m observações $Y = (Y_i, Y_{i+1}, \dots, Y_m)$ com n entradas X relacionadas à cada Y_m , ocorre da seguinte maneira: Matematicamente o neurônio é representado por um vetor global de n pesos sinápticos (W_0, W_1, \dots, W_n) , e um valor de bias b . O processo do algoritmo abaixo ajustará esses pesos e bias.

- 1) Coleta do valor supervisionado inicial Y_0 , que é a primeira observação histórica do conjunto de dados.
- 2) Coleta do vetor de variáveis $(X_{00}, X_{01}, \dots, X_{0n})$ relacionados à saída supervisionada Y_i .
- 3) Inicialização dos pesos sinápticos (W_0, W_1, \dots, W_n) com valores aleatoriamente pequenos.
- 4) Definição de um valor real entre 0 e 1 para a taxa de aprendizado α .
- 5) Definição do valor inicial de bias b .
- 6) Definição da função de ativação δ
- 7) Definição do limite de épocas de treino l_e , se não houver, definir l_e para -1.
- 8) LAÇO: Percorrer e atribuir em (Y_d) de Y_0 até Y_m , inicializando o contador de épocas $c_e = 0$, enquanto $CP(e)$ for verdadeiro E/OU (l_e) for menor que c_e :
- 9) Nova coleta do vetor de variáveis $(X_{d0}, X_{d1}, \dots, X_{dn})$ relacionados à saída supervisionada Y_d .
- 10) Incrementar 1 em c_e .
- 11) Cálculo do sinal $u = \sum_{i=1}^n X_{di} * W_i + b$
- 12) Aplicação de valor de U na função de ativação δ obtendo $Y_u = \delta(u)$
- 13) Obtenção do erro $e = Y_d - Y_u$
- 14) Aplicação do critério de parada CP em relação ao erro e ; $CP(e)$
- 15) SE $CP(e)$ for APROVADO, o algoritmo continua, SE REPROVADO, a aprendizagem é feita ajustando $b = b + (\alpha * (e))$, e ajustando o vetor $W_i = W_i + (\alpha * e * X_{di})$.

G. Rede Perceptrons Múltiplas Camadas - MLP

A solução base para se combinar 2 ou mais perceptrons a fim de se resolver um problema com a combinação de 2 ou mais soluções lineares, é a utilização de um perceptron combinador de sinal de saída, já que cada perceptron pode ter

múltiplas entradas e somente uma saída. Dessa forma as redes neurais vão formando colunas de perceptrons interconectados. Cada coluna é denominada uma camada da rede neural. A última camada deve ter o número de perceptrons correspondente ao número de saídas desejadas. Na figura abaixo, encontra-se uma rede neural com duas camadas intermediárias e 3 saídas.

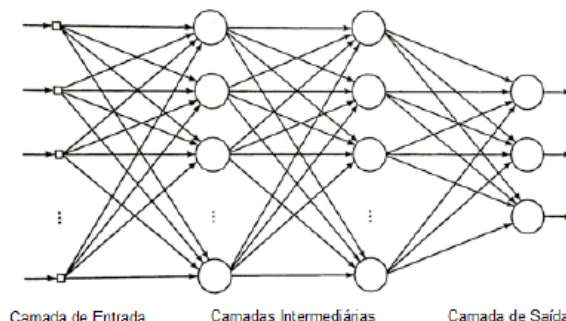


Figura 8. Rede de perceptrons com múltiplas camadas. Retirado de (ALMEIDA, 2013)

Esta rede denominada MLP (Multilayer Perceptron) possui uma camada de entrada onde cada nó representa uma variável a ser considerada ao problema a ser analisado, e pelo menos uma camada intermediária.

Nesta camada intermediária os neurônios possuem geralmente uma função de ativação sigmoidal logística ou tangente hiperbólica, e conceitualmente no mínimo 1 neurônio desta primeira camada oculta deve receber no mínimo 2 conexões de entrada. E uma camada de saída, na borda à direita, com o número de neurônios correspondente ao número de soluções buscadas.

a) *Treino e Validação da MLP:* O conjunto de dados de entrada na rede MLP supervisionada, deve ser dividido em 2 partes principais, Treinamento e Validação. É importante salientar que as observações de ambos os conjuntos devem originar do mesmo conjunto de dados para representar o mesmo problema, as observações Y_i relacionadas aos vetores i de variáveis $X_{i1}, X_{i2}, \dots, X_{ij}$ do conjunto de treino, devem ter a mesma estrutura, mesmo número j de variáveis X por vetor de entrada e devem representar o mesmo problema que as observações Y_i do conjunto de validação.

O conjunto de dados da validação e treino somados devem formar exatamente o conjunto original, sem informações excedentes ou em falta.

Os dados de treino e dados de validação, podem ser separados de forma aleatória, sendo os dados de treino os responsáveis pelos reajustes de pesos e capacidade de generalização da rede, e os dados de validação responsáveis pelo processo de validação pós-treino.

É importante adotar um bom critério de parada de treino, pois um treino prolongado tende a convergir em ajustes de pesos memorizados dos valores observados nos dados de treino e isso causa o fenômeno de overfitting na rede neural, que é a perda da capacidade de generalização. Pesos sinápticos que

sofrem processo prolongado de reajustes acaba "viciando" a rede neural para reconhecer apenas os dados de treino.

Parada do erro mínimo: O critério de parada do erro mínimo encerra o treinamento da RNA quando a mesma obtém um erro menor que o mínimo estipulado para o valor observado, este é o critério mais simples, adotado nos casos onde existe um limiar de erro já determinado pelo problema, entre o valor observado e o valor a se prever.

Parada por número de épocas: Um critério de parada pode ser limitado também ao número de épocas de treino. A determinação deste número de épocas pode ocorrer por tentativa e erro, visto que haverá a convergência de um número pequeno para uma baixa capacidade de aprendizado, e a convergência de um número grande para o processo de overfitting. Logo é necessário realizar experimentos que validem número de épocas fora desses intervalos de convergência.

Parada por validação cruzada: Por fim a validação cruzada é a técnica que se utiliza os dados dos conjuntos de validação e treino de forma cruzada, neste processo então, os dados de treino são utilizados no processo iterativo de aprendizagem, e no fim deste processo o conjunto é validado com os dados de validação, obtendo-se um novo erro de validação. A medição do erro de validação passa por um processo de avaliação em função do número de épocas, a fim de se detectar um ponto de número de épocas onde o erro quadrático médio da amostra de validação sofre uma curva de crescimento após um limiar de decréscimo.

É notório observar que o erro quadrático médio das amostras de treinamento sofrerá um decréscimo em função do aumento do número de épocas de treinamento, convergindo ao processo de overfitting, ou memorização da rede.

O ponto de parada no limite inferior do erro quadrático médio da amostra de validação, será o ponto ótimo de parada de treinamento.

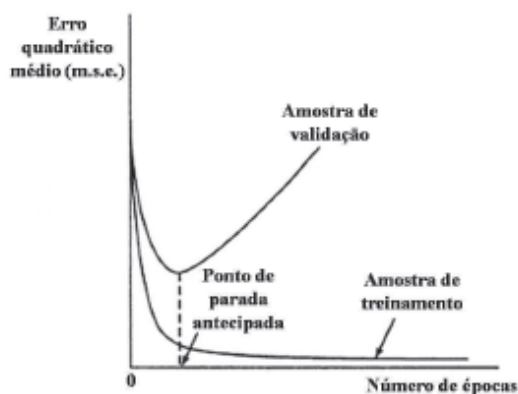


Figura 9. Ponto ótimo de parada da validação cruzada. Retirado de (MARIANO, 2014)

O treinamento de uma rede MLP para previsões de demanda de restaurantes universitários que obteve sucesso em (LOPES,

2008) e (ROCHA JOSÉ CELSO; MATOS, 2011) é a retro propagação de erro.

H. Perceptrons Múltiplas Camadas com Retro Propagação de Erro

A rede Perceptron Múltiplas Camadas com Retro Propagação de Erro (do inglês, M.L.P Backpropagation), faz o reajuste dos pesos sinápticos dos neurônios através de duas fases:

a) *Feed-forward*: Nesta primeira fase de treino os sinais $X_{i0}, X_{i1}, \dots, X_{in}$ com sua respectiva saída Y_i do conjunto de dados são apresentados à todos os neurônios da primeira camada. O processo de propagação do sinal de saída de cada neurônio segue o princípio do neurônio artificial apresentado na seção anterior, que envia o sinal de saída como um sinal de entrada ao neurônio seguinte.

b) *Feed-backward*: Nesta fase é obtido um valor de erro da camada de saída. Este erro é utilizado na equação de reajuste do peso sináptico das conexões dos neurônios da camada de saída com o sinal de saída dos neurônios da última camada oculta. E depois este erro é propagado realizando outra equação de reajuste de peso sináptico das conexões dos neurônios nas camadas anteriores, no sentido contrário em direção à camada de entrada. Isto permite que os pesos sinápticos de todas as camadas de intermediárias neste processo tenham seus pesos ajustados.

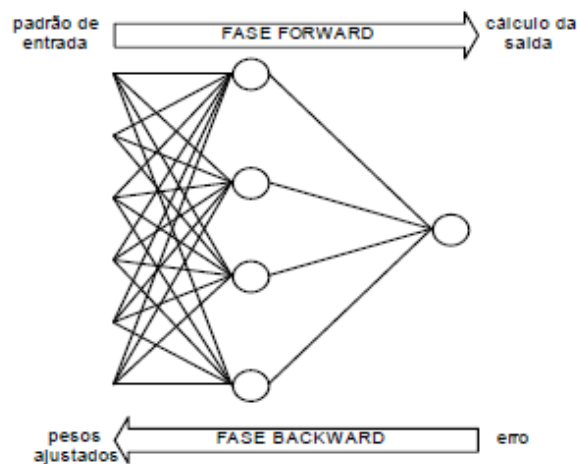


Figura 10. Fases de treino da MLP-Back-Propagation. Retirado de (ALMEIDA, 2013)

MLP Backpropagation em R.U em trabalhos relacionados: Em (LOPES, 2008) A rede neural perceptron de múltiplas camadas é utilizada para tratar a previsão de demanda do R.U da UFV, utilizando apenas como variáveis quantitativas as 5 ultimas observações anteriores ao dia a se analisar, e como variáveis qualitativas o dia da semana variando de segunda a sexta, em valores binários.

Em (ROCHA JOSÉ CELSO; MATOS, 2011) o trabalho realizado com neurônios artificiais para prever a demanda do R.U da UNESP, envolve apenas uma única camada de entrada e uma segunda camada para saída, porém com uma diversidade

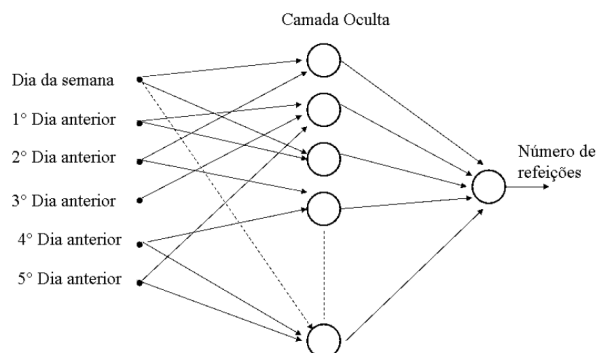


Figura 11. Rede Neural Perceptron de Múltiplas Camadas. Retirado de (LOPES, 2008).

maior de variáveis de entrada correlacionadas com o consumo do restaurante.

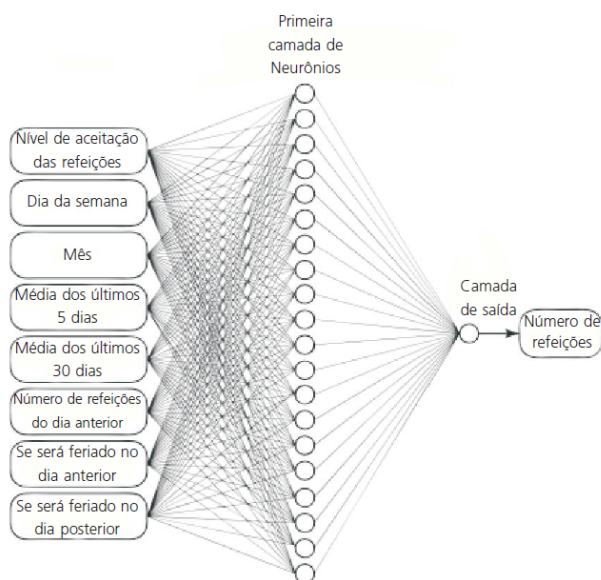


Figura 12. Rede Neural Perceptron de Múltiplas Camadas. Retirado de (ROCHA JOSÉ CELSO; MATOS, 2011)

Ambos os modelos possuem topologia de uma camada oculta para entrada dos dados, e uma camada de saída, que de acordo com (RAGA A. DE P.; CARVALHO, 2000) cita que através de uma análise de Cybenko, uma camada intermediária é o suficiente para aproximar qualquer função contínua e 2 camadas intermediárias são suficientes para aproximar qualquer função matemática, e devendo ser observado o fato de que em alguns casos, a utilização de 2 ou mais camadas pode facilitar o treinamento da rede, porém a utilização de um grande número de camadas intermediárias ou ocultas, é inviável, pois em cada uma delas a estimativa do erro se trata de uma estimativa da estimativa do erro da camada anterior, e este cascadeamento de estimativas pode se tornar menos preciso à medida que cresce.

(MARIANO, 2014) cita que o número de neurônios na primeira camada oculta é proporcional à dimensão do espaço de observação. Logo em modelos preditivos de demanda, supracitados, observa-se que no mínimo, na primeira camada oculta, é utilizado um número de neurônios igual ao número de variáveis que influenciam no dado preditivo.

Parâmetros de treino da MLP Backpropagation: A definição da função δ de ativação interfere na linearidade do modelo a ser analisado, sendo a função sigmoide a mais popular na literatura.

A taxa de aprendizagem, define a velocidade de reajuste dos pesos, podendo variar de 0 a 1. Ressalta-se que uma taxa próxima de 1 provoca picos oscilatórios na taxa de aprendizado, e taxa próxima de 0 provoca lentidão da convergência de aprendizagem. Valores comuns de utilização ficam entre 0,2 e 0,8.

TRABALHOS RELACIONADOS

I. ANÁLISES EM RESTAURANTES UNIVERSITÁRIOS

: No estudo estatístico feito por (LANDIM, 2016), foi analisada a correlação entre a temperatura e o consumo nos dias de vendas do restaurante universitário do campus ICT da Unifesp, sendo que os dados continham uma amostra das vendas do segundo semestre de 2016. Devido ao baixo volume de ocorrências, os dados foram submetidos à reamostragem via bootstrap. De acordo com os gráficos das amostras, identificou-se que a correlação mostrada nos gráficos da primeira metade do semestre e do período total do semestre formaram distribuições bimodais. Porém, na segunda metade do semestre formou-se uma distribuição unimodal. Portanto, concluiu-se que outras variáveis e outros modelos de análises podem ser utilizados para tal análise de demanda.

: (LOPES, 2008) faz o mesmo estudo deste cenário do ICT UNIFESP aplicado na Universidade Federal de Viçosa (UFV). Neste estudo, os dados utilizados foram somente o histórico de vendas do restaurante universitário, e nenhuma variável de ambiente foi coletada como temperatura, precipitação, número de alunos matriculados, etc. O algoritmo utilizado foi o Traincgp (Conjugate gradient backpropagation with Polak-Ribiere updates) no software Matlab. Este algoritmo não envolve o cálculo das derivadas segundas das variáveis e converge ao mínimo da função quadrática em um número finito de iterações como cita o autor. Foram então considerados para cada nó da rede neural, o dia da semana (como segunda, terça, quarta, quinta e sexta) e cada camada dessa rede utilizando os 5 dias anteriores para cada nó (as 5 segundas anteriores, 5 terças anteriores e assim sucessivamente) e, por fim, obtido um modelo pela rede que apresentou erro máximo de 3.

: (ROCHA JOSÉ CELSO; MATOS, 2011) também realiza o estudo de demanda no restaurante universitário da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), novamente com os métodos de redes neurais artificiais com backpropagation, e utilizando apenas como fonte de dados o histórico numérico das vendas realizadas, e outras variáveis intermediárias obtidas a partir deste como médias de

subconjunto de observações (médias de segundas-feiras), e a única variável de ambiente coletada foi o número de feriados próximos à observação de venda. No estudo do total de dias analisados, verifica-se que em 73% (187 dias), o método de média simples propiciou um maior erro em relação à RNA, que por sua vez ocasionou um erro maior nos 23% (69 dias) restantes. Em se tratando de menor desperdício, observa-se que a RNA apresenta erros maiores que 50 refeições em 13 dias, enquanto o método da média simples apresenta erros maiores que 50 refeições em 58 dias, concluindo-se então que o método de RNA foi bem mais eficiente do que o cálculo de média simples utilizado pela administração do restaurante universitário.

J. ANÁLISES EM OUTROS CENÁRIOS

: (RUAS, 2012) fazem uma análise de previsão de demanda de energia elétrica no estado do Paraná, entre os anos de 2004 e 2006, utilizando redes neurais artificiais e máquinas de vetores de suporte. Apesar de não ser o mesmo exemplo do cenário do restaurante universitário do ICT UNIFESP, temos na distribuição dos dados de consumo coletados como uma série temporal, e foi utilizado nesta pesquisa uma rede parcialmente recorrente de Elman, que permite a previsão de um passo de tempo à frente. Para que seja possível realizar a previsão vários pontos à frente, é necessário utilizar os valores já previstos, ou seja, a saída da rede, como entradas da mesma. Logo, como o cenário da demanda do RU no ICT UNIFESP se apresenta como uma série temporal e tentaremos prever a demanda um passo à frente, será interessante a este trabalho utilizar e comparar as técnicas previstas nesta referência.

: (ALMEIDA, 2013) analisa um cenário semelhante de demanda de energia elétrica, porém utilizando-se técnicas de previsão de demanda com Rede Neural Artificial do tipo Multilayer Perceptron combinado com lógica fuzzy que permite colocar variáveis de temperatura (entre outras) em um conjunto de regras que impactam no problema. Esta análise é interessante pois é semelhante à referência 2 que faz um estudo com redes neurais perceptron no mesmo tema deste trabalho de conclusão de curso, mas com a adição da lógica fuzzy que permite utilizar outras variáveis de ambiente impactantes, que não sejam somente o histórico numérico do consumo.

: (SILVA, 2010) também aplica técnicas de redes neurais para previsão de demanda de energia elétrica, com o estudo de variáveis climáticas, porém através de um modelo de MAPA SOM - (Self-Organizing Map) que é um tipo de rede neural desenvolvido para reconhecimento de padrões. Apesar de ser um modelo não supervisionado, o modelo é ideal para organizar as principais variáveis impactantes e descartáveis na previsão. O mapa som utilizado pelo autor apresenta os dados associados aos seus neurônios de forma que padrões similares encontram-se em neurônios contíguos, tendo uma organização topológica. Deste modo é possível se extrair relações abstratas entre as variáveis do vetor de dados através da sua posição nos mapas componentes, que por meio de uma escala de cores mostram a quantidade de uma variável específica em cada neurônio do mapa.

K. TESE COMPARATIVA DE DIVERSOS MÉTODOS DE PREVISÃO DE DEMANDA.

: (JUNIOR, 2007) realiza um trabalho de comparação entre diversas técnicas de Métodos estocásticos e Rede Neural Artificial (RNA), para a previsão da demanda de produtos cosméticos, distribuídos em séries temporais. Entre eles Redes Neurais tipo feedforward, com o algoritmo backpropagation que foi o principal foco no trabalho de previsão de RU na Universidade Federal de Viçosa e na Universidade Estadual Paulista Júlio de Mesquita Filho. Será realizado a análise dessas comparações e seleção de principais algoritmos que possam ser aplicados e comparados neste trabalho.

PROTOCOLO EXPERIMENTAL

L. Base de dados

Foi Estruturado o conjunto de dados do R.U do ano de 2017 como base de treino conforme e os dados de 2018 como base de teste. A medição de qualidade do modelo será feita com dados de 2018. A partir da data de cada dia, será coletado o valor categórico deste dia sendo segunda, terça, quarta, quinta ou sexta feira, e as vendas de até 5 dias anteriores à data que se deseja prever, fundamentado nos trabalhos de (LOPES, 2008) e (ROCHA JOSÉ CELSO; MATOS, 2011)

Como hipótese deste trabalho, será adicionado o valor categórico das estações do ano, e distância em dias da observação posterior e anterior. Dado que distâncias grandes simbolizam recessos, feriados prolongados, greves ou férias, que impactam diretamente na frequência de alunos na universidade e consequentemente no restaurante.

Também serão analisadas variáveis climáticas juntos com os dados de consumo, de forma a analisar a influência de fatores externos como temperatura média ambiente e precipitação. Tais dados podem ser obtidos de forma gratuita pelo BDMEP - Banco de Dados Meteorológicos para Ensino e Pesquisa, pertencente à instituição pública INMET - Instituto Nacional de Meteorologia, pertencente ao MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO do Governo Brasileiro. É necessário um cadastro no site <http://www.inmet.gov.br/portal/index.php?r=bdmep/bdmeppara> a obtenção dos dados.

M. Modelo de RNA Proposto

EXPERIMENTOS

N. Ferramentas Utilizadas

Os códigos e experimentos a seguir foram realizados com a biblioteca Keras operada através do Jupyter Notebook, e alguns experimentos foram também validados com a ferramenta (em Java) Weka.

O. Importação das bibliotecas

No Jupiter Notebook foram importadas as seguintes bibliotecas:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

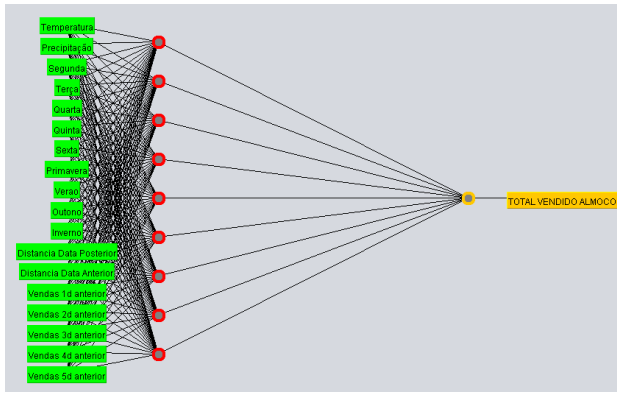


Figura 13. Parâmetros e modelo de RNA proposto para previsão do consumo no restaurante do ICT.

```
import seaborn as sns
from sklearn.preprocessing import
    LabelEncoder, OneHotEncoder
from sklearn.model_selection import
    train_test_split
from keras.models import Sequential
from keras.layers import Dense
from keras.wrappers.scikit_learn import
    KerasRegressor
from keras.callbacks import ModelCheckpoint
from sklearn import metrics
import os
```

P. Importação dos dados

Os dados de 2017 para treino, contidos em formato .csv foram importados para o conjunto X-train contendo o valor de todos os atributos de cada observação por data e Y-train contendo o valor total de vendas de almoço por data. O mesmo foi realizado para os dados de 2017. Os atributos de entrada de treino ficaram originalmente então no seguinte formato:

```
original_data.columns
Index(['Data', 'Temperatura',
      'Precipitacao', 'Segunda', 'Terca',
      'Quarta',
      'Quinta', 'Sexta', 'Primavera', 'Verao',
      'Outono', 'Inverno',
      'Distancia Data Posterior', 'Distancia
      Data Anterior',
      'Vendas 1d anterior', 'Vendas 2d
      anterior', 'Vendas 3d anterior',
      'Vendas 4d anterior', 'Vendas 5d
      anterior', 'Total'],
      dtype='object')

X_train = original_data.iloc[:,1:19]
X_train.values
array([[ 29.9 ,  77.25,  0. , ..., 101. ,
        178. , 215. ],
       [ 29.2 ,  74.25,  1. , ..., 178. , 215. ,
        263. ],
       [ 31.9 ,  67.5 ,  0. , ..., 215. , 263. ,
        145. ],
       ...,
       [ 24.1 ,  88.25,  1. , ...,  0. ,  0. ,  0. ],
```

```
[ 20.3 ,  78.75,  0. , ...,  0. ,  0. ,  0. ],
[ 25.2 ,  83.5 ,  0. , ...,  0. ,  0. ,  0.
]])
```

```
y_train = original_data.iloc[:,19].values
y_train
array([ 42,  70,  64, 101, 178, 215, 263,
       145, 295, 351, 387, 294, 194,
       460, 363, 437, 314, 212, 396, 336, 427,
       312, 218, 407, 428, 306,
       220, 367, 352, 463, 320, 306, 451, 353,
       95, 149, 209, 227, 202,
       233, 391, 358, 456, 335, 323, 469, 319,
       297, 411, 370, 412, 331,
       255, 446, 394, 472, 361, 263, 428, 336,
       470, 343, 232, 449, 386,
       477, 377, 347, 480, 359, 224, 450, 375,
       491, 347, 266, 415, 405,
       464, 358, 253, 446, 388, 523, 346, 245,
       438, 361, 503, 319, 39,
       45, 97, 115, 138, 134, 211, 308, 353,
       401, 382, 321, 352, 428,
       453, 433, 376, 449, 403, 286, 408, 393,
       442, 431, 17, 12, 17,
       27, 16, 141, 117, 42, 20, 363, 162, 35,
       351, 21, 23, 27,
       28, 162, 23, 14, 84, 41, 9, 15, 15, 17,
       26, 10, 14,
       17, 26, 10, 22, 14], dtype=int64)
```

Da mesma forma os atributos de teste foram importados e configurados.

Q. Análise Exploratória dos dados

Para uma melhor compreensão do domínio dos atributos, foram realizadas algumas análises nos dados de treino, onde se observou uma frequência maior de dias com precipitação acima de 70%, temperatura média de 25 graus, os dias mais predominantes da semana foram segundas e terças, e uma venda média de 266 refeições (almoço por dia).

R. Normalização dos dados

Para realizar o treino, foi necessário normalizar os atributos de treino e teste, através da biblioteca StandardScaler

```
from sklearn.preprocessing import
    StandardScaler
scaler = StandardScaler().fit(X_train)
standardized_X = scaler.transform(X_train)
standardized_X_test =
    scaler.transform(X_test)
```

Os atributos de entrada ficaram então no seguinte formato:

```
array([[ 1.13495156,  0.30172306,
        -0.52522573, ..., -1.02827871,
        -0.53945104, -0.29949702],
       [ 0.96953093,  0.06398563,  1.90394328,
        ..., -0.54679855,
        -0.30940401, -0.00345878], ...])
```

S. Construção dos Modelos de RNA

3 Modelos então foram construídos através dos seguintes códigos:

```
model = Sequential()
model.add(Dense(standardized_X.shape[1],
activation='relu',
input_dim=standardized_X.shape[1]))
model.add(Dense(1))
model.compile(optimizer='adam',
loss='mse',
metrics=['mae'])
```

Neste modelo com o número de entradas na primeira camada igual ao número de atributos, e 1 neurônio de saída, o melhor desempenho foi adotado com o otimizador "adam", função "relu", tamanho do batch no mesmo tamanho de amostras 148. Quando a função "adam" foi trocada para "rmsprop" e "relu" trocada para "sigmoid" houve perda de desempenho significativa. As medidas máximas de MSE com a função sigmoid e otimizador rmsprop foram de 72.5417.

O número de épocas foi definido para 5000, porém foi realizado a parada antecipada por validação cruzada sobre os dados de validação.

```
model.fit(standardized_X,
y_train,
batch_size=36,
epochs=5000,
validation_data=(standardized_X_test,y_test),
callbacks=[early_stopping_monitor])
```

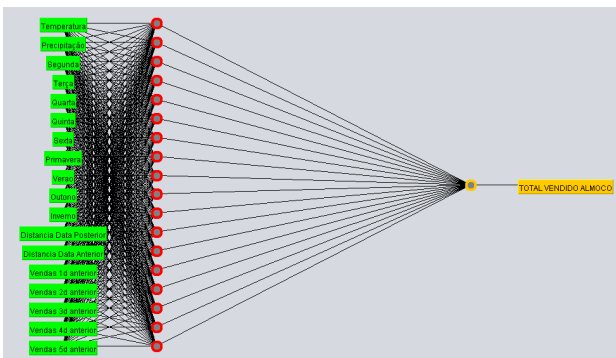


Figura 14. Modelo 1 inicial com número de neurônios igual a número de atributos, sem camadas ocultas.

Este modelo obteve as seguintes métricas: mean-absolute-error: 42.7853 - val-loss: 8063.6133 - val-mean-absolute-error: 60.9182

Como segundo modelo de previsão, foi construído uma rede com 9 neurônios de entrada, a metade dos atributos e 1 de saída, sem camadas ocultas.

```
model1 = Sequential()
model1.add(Dense(9,
activation='relu',
input_dim=standardized_X.shape[1]))
model1.add(Dense(1))
```

```
model1.compile(optimizer='adam',
loss='mse',
metrics=['mae'])
```

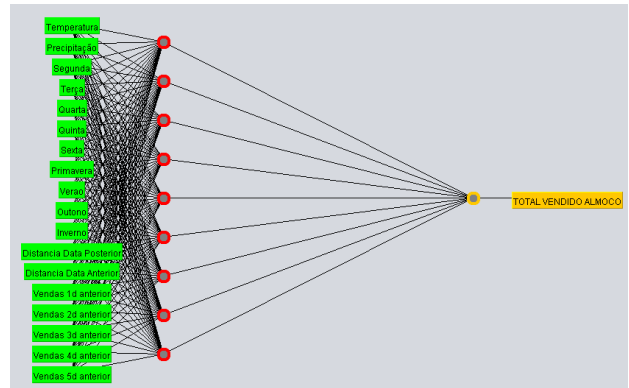


Figura 15. Modelo 2 com número de neurônios reduzido à metade do número de atributos, sem camadas ocultas.

Este modelo obteve as seguintes métricas: mean-absolute-error: 60.3270 - val-loss: 17238.3887 - val-mean-absolute-error: 95.9984

Um terceiro modelo foi construído com 2 camadas ocultas, que obteve o melhor resultado entre os anteriores:

```
model3 = Sequential()
model3.add(Dense(standardized_X.shape[1],
activation='relu',
input_dim=standardized_X.shape[1]))

model3.add(Dense(9,activation='relu',
input_dim=standardized_X.shape[1]))

model3.add(Dense(3,activation='relu',
input_dim=9))

model3.add(Dense(1))

model3.compile(optimizer='adam',
loss='mse',
metrics=['mae'])
```

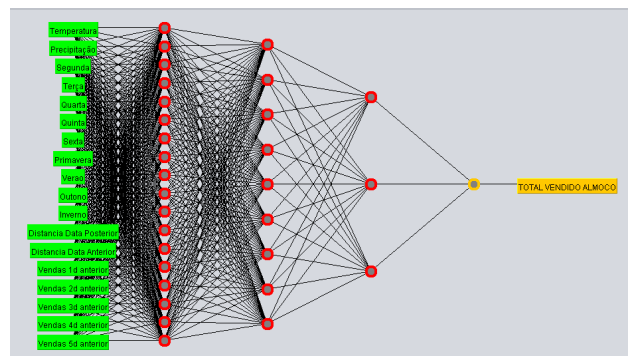


Figura 16. Modelo 3 com número de neurônios na camada de entrada igual a número de atributos, adição de 2 camadas ocultas.

Este modelo obteve as seguintes métricas: mean-absolute-error: 43.0977 - val-loss: 7648.9710 - val-mean-absolute-error: 59.6000

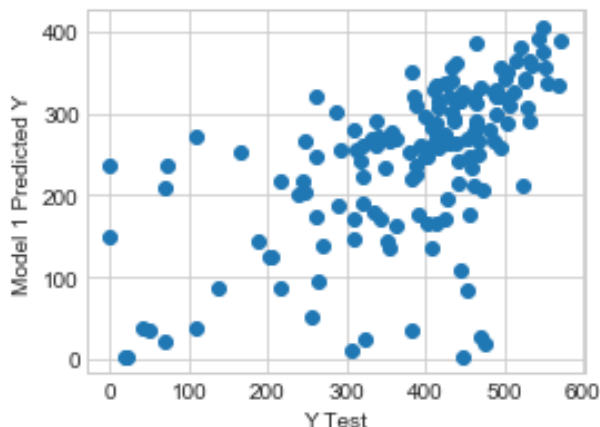


Figura 17. Plotagem e distribuição dos dados previstos e dados reais do modelo 1.

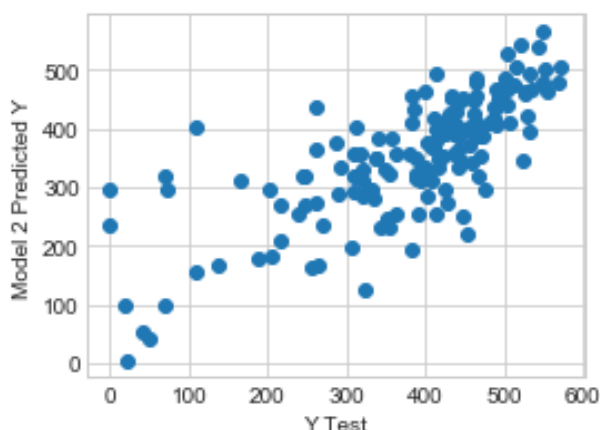


Figura 18. Plotagem e distribuição dos dados previstos e dados reais do modelo 2.

CONCLUSÃO

Como o objetivo do trabalho é a minimização do desperdício de alimentos, que hoje atua entre 30 a 50% das refeições produzidas, o melhor modelo treinado não atinge o % desejado de acerto. No caso o modelo 3.

No entanto ao comparar a tendencia linear do mesmo dos valores previstos com os reais, observa-se que este segue o comportamento de frequência dos alunos principalmente nas alternâncias de semestre.

Os dados de treino também estão escassos 148 observações em 2017 que estão incompletas iniciando-se em abril e 150 observações em 2018 finalizando em 31/10/2018, devido o novo sistema de refeições e seu banco de dados ter sido implementado recentemente. Logo concluí-se que o sucesso do treino pode ser obtido tendo mais anos completos de dados

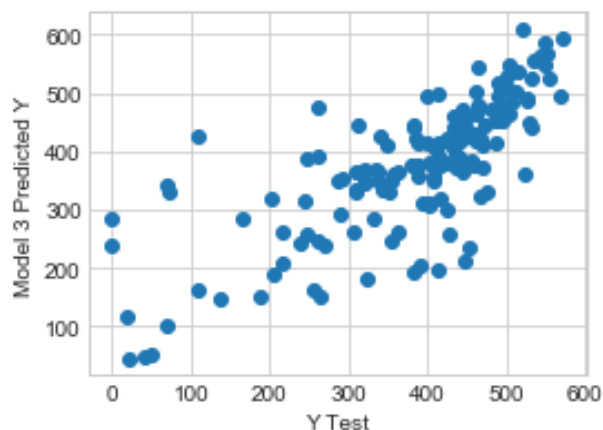


Figura 19. Plotagem e distribuição dos dados previstos e dados reais do modelo 3.

para treino e no mínimo 1 ano para o processo de validação cruzada.

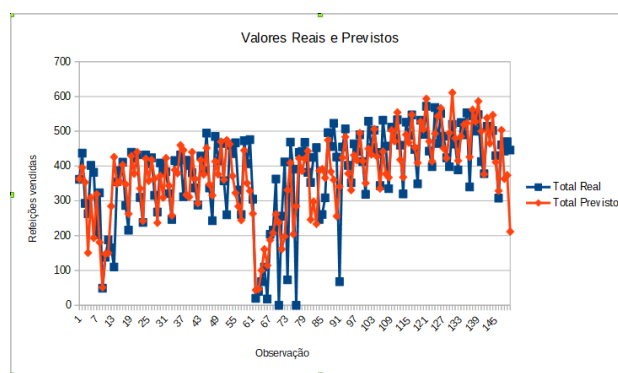


Figura 20. Total Real VS Total Previsto

REFERÊNCIAS

- ALMEIDA, V. A. d. *Previsão De Carga Através De Modelos Neuro-fuzzy*. 91 p. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, 2013. Disponível em: <<http://monografias.poli.ufrj.br/monografias/monopoli10007620.pdf>>. Citado 5 vezes nas páginas page.22, page.55, page.66, page.77 e page.99.
- BRYAN KOLB; ARIF, M. Harnessing the power of neuroplasticity for intervention. In: *FRONTIERS IN HUMAN NEUROSCIENCE*. [S.l.: s.n.], 2014. p. 13. Citado na página page.44.
- HAYKIN, S. *Neural networks, a comprehensive foundation*. [S.l.]: Macmillan, 1994. Citado na página page.22.
- JUNIOR, A. M. *Análise de métodos de previsão de demanda baseados em séries temporais em uma empresa do setor de perfumes e cosméticos*. Dissertação (Mestrado) — Pontifícia Universidade do Paraná. Curitiba, 2007. Citado 4 vezes nas páginas page.22, page.33, page.55 e page.99.
- LANDIM, D. D. *Análise de demanda no restaurante universitário do instituto de ciência e tecnologia da unifesp - universidade federal de são paulo*. *Projeto da Disciplina de Simulação de Sistemas*, Universidade Federal de São Paulo, v. 1, n. 1, p. 6, 2016. Disponível em: <<https://github.com/ddlandim/monografy-ann-demand-prediction/tree/master/referencias>>. Citado na página page.88.
- LOPES, L. Modelos para previsão de demanda no restaurante universitário utilizando técnicas de redes neurais. *Congresso Nacional De Excelência Em Gestão*. Universidade Federal De Viçosa., Universidade Federal De Viçosa, IV, n. 1, p. 14, 2008. Disponível em: <<http://>>

//www.novos cursos.ufv.br/projetos/ufv/gesep/www/wp-content/uploads/5.>

Citado 4 vezes nas páginas page.22, page.77, page.88 e page.99.

LUGER, G. F. *Inteligência Artificial. Estruturas e Estratégias para a Solução de Problemas Complexos*, 4ª ed. [S.l.]: Porto Alegre: Bookman, 2004. Citado na página page.33.

MARIANO, F. C. M. Q. *REdes Neurais na predição de valores energéticos de alimentos para aves*. 130 p. Tese (Doutorado) — Universidade Federal de Lavras, 2014. Citado 4 vezes nas páginas page.22, page.55, page.77 e page.88.

RAGA A. DE P.; CARVALHO, A. P. L. F. L. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: Livros técnicos e científicos, 2000. Citado na página page.88.

ROCHA JOSÉ CELSO; MATOS, F. D. F. Utilização de redes neurais artificiais para a determinação do número de refeições diárias de um restaurante universitário. *Rev. Nutr., Campinas*, Universidade Estadual Paulista Júlio de Mesquita Filho., n. 1, p. 8, 2011. Citado 4 vezes nas páginas page.22, page.77, page.88 e page.99.

RUAS, G. I. e. S. Previsão de demanda de energia elétrica utilizando redes neurais artificiais e support vector regression. *Copel Distribuição S/A (COPEL)*, Universidade Federal do Paraná(UFPR), v. 1, n. 1, p. 10, 2012. Disponível em: <http://www.dcc.fc.up.pt/~ines/enia07_html/pdf/27927.pdf>. Citado 2 vezes nas páginas page.22 e page.99.

SILVA, C. e. a. Metodologia para previsão de demanda de energia elétrica em curtíssimo prazo baseada em um modelo probabilístico discreto. *Desenvolvimento de Algoritmos e Software para Previsão de Carga na Área de Concessão da MuxEnergia - Horizonte de Curtíssimo Prazo*, Universidade Federal do Pampa - UNIPAMPA, v. 1, n. 1, p. 4, 2010. Disponível em: <http://alegrete.unipampa.edu.br/gesep/wp-content/uploads/2010/05/Previs%C3%A3o-de-demanda-utilizando-m%C3%A9todo-de-Markov_Final.pdf>. Citado 2 vezes nas páginas page.22 e page.99.

ZORZETTO, R. Recontagem de neurônios põe em xeque ideias da neurociência. *Revista Fapesp* 192, Fapesp, v. 1, n. 1, p. 6, 2012. Citado na página page.44.