|  |
| --- |
| **Estudo do algoritmo de aprendizado de máquina *Support Vector Regression* (SVR)como ferramenta de apoio na gestão de refeitórios**  **Autores (as):**  **Igor Martins Ferreira**  **Lucas Silva Sousa**  **Renan Dias de Oliveira**  **Orientador (a):**  **Gabriel de Lara Baptista** |

**Resumo**

Grande parte das decisões de um gestor de refeitórios ainda ocorrem de maneira empírica ao estimar o quanto se deve produzir de refeições. Considerando esta premissa, o presente artigo relata como foi o estudo e a implementação de um algoritmo de aprendizado de máquina com o objetivo de auxiliar na tomada de decisão mais assertiva. O estudo tomou como base os dados coletados de um cenário específico no ano de 2018, dentro de uma instituição de ensino, além de algumas variáveis que aparentavam influenciar na presença das pessoas em seu refeitório. Na etapa onde houve análise das correlações dos dados coletados, observou-se que tais dados não tinham correlações lineares, o que não impediria de correlacioná-las de forma não linear. Dessa forma utilizou-se o *Support Vector Regression (SVR),* um algoritmo com alta capacidade de generalização, para estimar o número de refeições ideal. Mesmo assim, os resultados obtidos não se mostraram satisfatórios, tendo em vista que houve um grande erro absoluto e um desvio padrão alto no número de refeições estimadas.

**Palavras-chave**: Aprendizado de máquina, *SVR*, gestão de refeitórios, desperdício de alimentos.

**Abstract**

*The restaurant manager's decision still occurs empirically when estimating how many meals need to be produced. Considering this premise, this article reports the study and implementation of a machine learning algorithm in order to assist in more assertive decision making. The study was based on data collected from a specific scenario in 2018, within a teaching institution, in addition to some variables that appeared to influence the presence of people in its restaurant. In the stage of the correlations analysis of the collected data, was observed that such data did not have linear correlations, which would not prevent them from being correlated non-linearly. Thus, the Support Vector Regression (SVR), an algorithm with high generalization capacity, was used to estimate the ideal number of meals. Even so, the results obtained were not satisfactory, considering that there was a large absolute error and a high standard deviation in the number of estimated meals.*

**Keywords**: *Machine Learning*, *SVR, restaurant management, food waste.*

1. **Introdução**

Uma grande parcela da comida produzida no mundo é jogada fora sem ter a chance de ser ingerida. A Organização das Nações Unidas para a Alimentação e a Agricultura (FAO), alertou que, anualmente são desperdiçadas ou se perdem ao longo das cadeias produtivas cerca de 1,3 bilhão de toneladas de comida. O órgão ainda afirma que este volume representa 30% da comida produzida por ano no planeta, comida essa que poderia ser destinada às 821 milhões de pessoas que ainda passam fome no mundo (FAO, 2018).

Segundo a FAO (2018) o desperdício e perda estão presentes em todos os locais que produzem, vendem ou transportam alimentos, e até mesmo na mesa do consumidor, sendo que este último grupo representa 28% do desperdício. Dentro deste grupo estão as Unidades de Alimentação e Nutrição (UAN’s), que são entendidos como locais de produção e distribuição e alimentação para coletividades, tais como refeitórios escolares, restaurantes, refeitórios industriais e praças de alimentação (TEIXEIRA, 2007).

Como é dito por Júnior e Teixeira (2010) e Abreu et al (2012), a quantidade de sobras está relacionada com o planejamento inadequado de refeições e o número de pessoas que frequentam o estabelecimento. Por isso, é importante a realização de um planejamento com antecedência, usando registros das quantidades produzidas anteriormente para adequar a produção futura.

Com esses números em mente, não se pode ignorar o fato de que se estes alimentos puderem ser melhor aproveitados, mais pessoas teriam acesso aos mesmos, reduzindo assim a quantidade de pessoas que passam fome no mundo, mas seria possível utilizar a tecnologia para auxiliar esse processo que hoje ocorre de forma empírica, ou seja, pela experiência da pessoa que irá estimar quanto de comida produzir?

Dentro do cenário de evolução tecnológica, procura-se associar a tecnologia às técnicas existentes visando uma melhoria de desempenho destes métodos. Sendo assim o presente trabalho pretende estudar o tema Inteligência Artificial com o intuito de auxiliar o gestor da cozinha na escolha da quantidade de comida ou número de refeições que devem ser produzidas para um determinado dia da semana através de técnicas de aprendizado de máquina que utilizarão dados do passado para prever o futuro.

1. **Referencial Teórico**
   1. **Lixo alimentar**

Dentro da literatura, existem diferentes definições para o lixo alimentar. Por esta razão, esse capítulo pretende descrever os conceitos que serão utilizados como referência a esse assunto ao longo do trabalho.

Basicamente, o lixo alimentar tem dois tipos principais, o desperdício e a perda. Essa última é definida como a diminuição involuntária da disponibilidade do alimento, seja na produção, no armazenamento ou transporte dos mesmos (FAO, 2020).

Já o desperdício é resultado da decisão voluntária de se jogar fora a comida, depois de preparada e apta para consumo (Peixoto & Pinto, 2016). O desperdício de alimentos pode ser dividido entre dois subgrupos distintos para ajudar a identificar melhor o problema com base na sua causa, são eles:

Sobra Limpa: É a comida que preparada, porém não consumida, ou seja, são os alimentos que ficam na cuba de alimentos (COLOG, 2018).

Sobra Suja: é a sobra de alimento no prato do comensal, também conhecido como resto ingesta (SCOTTON, 2010).

O presente trabalho trata apenas da sobra limpa, pois ela está diretamente ligada a quantidade de comida que o gestor da UAN decidiu produzir para determinado dia, segundo Pereira (2018) quase todos os serviços de refeitórios e restaurantes se baseiam na intuição dos gestores para estimar a quantidade de comida que deve ser produzida, esse método de estimativa pode levar a superestimação do número de refeições que devem ser produzidas gerando desperdício na forma de sobra limpa.

* 1. **Inteligência Artificial**

A inteligência artificial é uma vertente da ciência da computação na qual se realiza o estudo das faculdades mentais através de modelos computacionais e de cálculos que possibilitam perceber, raciocinar e agir (CHARNIAK e MCDERMOTT, 1986; WINSTON, 1992).

A palavra inteligência é proveniente do latim *inter* (entre) e *legere* (escolher), ou seja, é aquilo que possibilita escolhas sendo elas racionais ou não, é a habilidade de corporificar determinada tarefa de forma eficiente, já a palavra artificial originou-se do latim *artificiale*, reflete algo não natural, feito pelo homem. (FERNANDES, 2005).

Dessa forma, é possível dizer que inteligência artificial é o esforço para fazer com que a ciência da computação desenvolva computadores e modelos para pensarem, de maneira semelhante ao funcionamento da mente humana no sentido completo e/ou literal (HAUGELAND, 1985), para exercer atividades como tomada de decisão, resolução de problemas e de aprendizado (HELLMAN, 1978).

Uma das vertentes existentes na inteligência artificial é o aprendizado de máquina, que lida com a implementação de programas de computador que podem aprender de maneira autônoma (HOSCH, 2020) e executar a tarefa aprendida. Esses programas são subdivididos de acordo com o tipo de tarefa executada:

* Classificação: É um algoritmo que tenta aprender a relação entre um grupo de variáveis de base e a variável alvo, com a finalidade de quando receber novos dados de base, acerte qual é o alvo automaticamente, pois já está treinado para isso. (AGGARWAL, 2014)
* Regressão: É um método de investigar as relações funcionais entre variáveis de base e variáveis alvo, mas neste caso a ligação entre elas é matemática, feito por uma equação, e visa gerar um valor numérico na variável alvo, ao invés de uma classe. (CHATTERJEE, 2015)

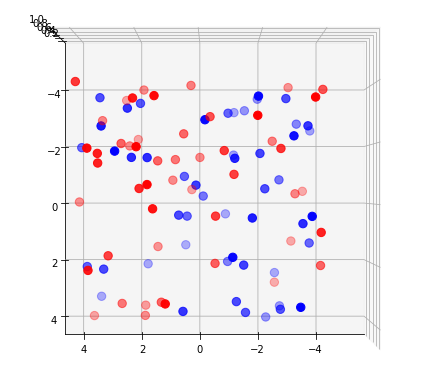
Os programas também são divididos entre dois tipos. Supervisionados e não Supervisionados.

Algoritmos supervisionados utilizam resultados já conhecidos para aprender, enquanto não supervisionados aprendem de maneira autônoma de acordo com os dados informados. Cada algoritmo desempenha as tarefas acima de uma maneira diferente, entre eles podemos citar as *Support Vector Machines*, que podem ser usadas para tarefas de classificação e regressão (SENAI, 2020).

* + 1. ***Support Vector Machines***

*Support Vector Machines (SVM)* é uma técnica de aprendizado de máquina utilizada para criação de um classificador, utilizado para separar e categorizar grupos de entradas de dados em uma vou mais categorias (PAULA, 2016), comumente muito útil nos casos em que os dados não são linearmente separáveis. Na figura 1 temos duas classes diferentes, os pontos vermelhos e azuis. Um *SVM* adiciona um hiperplano que separa os pontos vermelhos dos azuis pela máxima distância possível entre eles. O hiperplano em geometria é uma generalização do plano em N–1 dimensões, na figura 1 os dados estão plotados nos eixos X e Y, ou seja, duas dimensões. Neste caso o hiperplano traçado pelo algoritmo é uma reta, porém ela não é capaz de separar os pontos, desse modo podemos dizer eles não são linearmente separáveis.

Figura - Ilustração do SVM



Fonte: Autoria própria

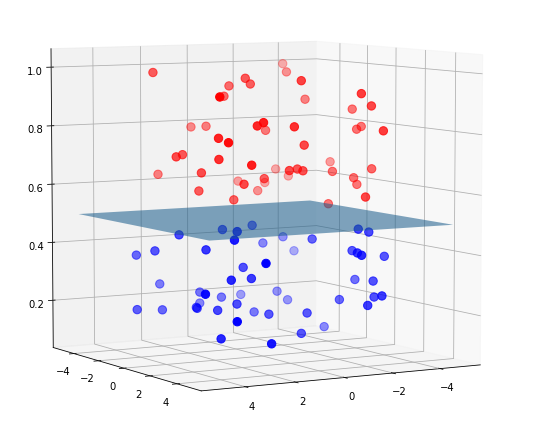
Para contornar esse problema, um modelo de *Support Vector Machine* organiza e esquematiza os dados de entrada em um espaço e cria um hiperplano para separar os dados como pode-se notar na figura 2.

Figura - Ilustração do hiperplano, própria autoria

* + 1. ***Support Vector Regression***

*Support Vector Regression* (SVR) é uma derivação do *SVM* que é um tipo de algoritmo utilizado para solução de problemas de regressão, isto é, prever valores de variáveis contínuas de uma função com múltiplas variáveis. Trata-se de um método de aprendizado de máquina supervisionado, ou seja, utiliza entradas de dados com resultados conhecidos para treinamento (AWAD e KHANNA, 2015). Uma maneira simples de entender o funcionamento dele é através de uma reta como exemplificado na figura 3. O épsilon é um dos hiperparâmetros para ajustar o modelo, ele é basicamente define um limite de tolerância, todos os dados que estiverem dentro da função da reta Ax + b + Epsilon, não são considerados como erro.

Uma imagem contendo mapa, pipa, texto, diferente

Descrição gerada automaticamente

Figura - Ilustração de funcionamento do SVR, própria autoria.

O SVR permite ainda a utilização de alguns métodos para refinar, melhorar ou entender a performance do algoritmo, entre eles pode-se citar o *Grid Search* e o *Permutation Feature Importance*.

* ***Grid Search:*** O SVR possui um conjunto de hiperparâmetros que tem valores numéricos definidos pelo usuário e que devem ser ajustados para que o algoritmo obtenha melhor resultados em sua precisão. O *Grid Search* é implementado escolhendo um intervalo valores para estes hiperparâmetros serem testados, eles são combinados entre si através de múltiplas iterações de treinamento com diferentes valores até que todas as possibilidades sejam avaliadas. A cada iteração o resultado de performance do algoritmo é calculada e registrado. Por fim, o melhor conjunto de hiperparâmetros é exibido em uma lista, para que se possa escolher qual deve ser aplicado no modelo (PEDREGOSA *et al.* 2011).
* ***Permutation Feature Importance***: Trata-se de uma técnica utilizada para descobrir em quais variáveis que o modelo mais se baseia e o quanto sua variação impacta na hora de gerar os resultados (PEDREGOSA *et al.* 2011).

O que levou o SVR à ser escolhido para ser aplicado neste caso é que de acordo com Awad e Khanna (2015), o SVR possui alta acurácia em suas previsões e alta capacidade de generalização, ou seja, descrever ou capturar os relacionamentos casuais não lineares entre as variáveis, ou seja, descobrir padrões não vistos anteriormente.

1. **Metodologia**

Considerando o objetivo da pesquisa, o trabalho passou por uma etapa inicial de revisão bibliográfica que sugeriu como técnica de aprendizado de máquina a ser utilizada para o problema algoritmos de regressão. Destes, o que mais aparentou ser útil para a pesquisa é o *Support Vector Regression.*

Se tratando de um problema de regressão, isto é, olhar para os dados do passado e tentar com base nisso prever o futuro, para que fosse possível atacar diretamente o problema do desperdício, seria necessário ter acesso à quantidade de comida produzida e a quantidade desperdiçada. No entanto, não foi possível obter tais informações. Desta forma, o algoritmo focou em gerar uma previsão de quantas pessoas irão à UAN com base nos dados históricos de refeições servidas, combinados com outros dados relevantes. A intenção é que essa estimativa auxilie a gestão do refeitório a planejar com antecedência o número de refeições que deverão ser produzidas naquele dia, para que assim o desperdício por sobras limpas diminua.

Para testar essa hipótese, o refeitório da Faculdade de Tecnologia Termomecânica foi escolhido como piloto para os testes. O refeitório serve as refeições no horário do almoço das 11:00 às 13:00 horas, e a janta das 18:15 até às 19:00 horas. A maior parte das refeições são servidas no período da manhã, onde estão presentes alunos e colaboradores em maior quantidade. O fluxo de alunos é mais ou menos constante na Instituição, já que uma das regras de aprovação nos cursos exige no mínimo 75% de presença de cada aluno nas aulas, fator esse que pode ter impacto na quantidade de pessoas que se alimentam na faculdade. Os alunos também são comunicados dos cardápios com antecedência, geralmente no começo da semana.

O estudo se baseou nos dados do jantar, visto que a variação das refeições é maior a noite, pois há menos colaboradores se alimentando todos os dias.

* 1. **Coleta de dados**

Primeiramente foram coletados alguns dados junto a faculdade. Esses dados foram organizados em forma de tabela com cada coluna da tabela correspondendo a uma variável dependente ou independente que foram utilizadas no treinamento do algoritmo de previsão.

A escolha das variáveis independentes foi baseada em estudos anteriores e outros fatores relativos ao ambiente da faculdade também foram levados em consideração pelo seu potencial de influência na demanda por comida.

É possível observar uma introdução aos dados coletados na Tabela 1, juntamente com uma breve descrição das variáveis.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variável** | **Tipo** | **Descrição** | **Fonte** |
| Dia da semana | Valores 1 - 5 | Segunda, Terça, Quarta, Quinta e Sexta. | Calendário |
| Dia do bimestre letivo | 1 - n | Do primeiro dia do bimestre até a última avaliação do bimestre. | Calendário FTT |
| Precipitação | Escala em mm | Quantidade de chuva registrada para o determinado dia. | Instituto Nacional de Meteorologia |
| Refeições na janta | Valor numérico | Quantidade de pessoas ou, se preferir, refeições servidas. | UAN FTT |
| Guarnição | Variável categórica | Principal guarnição servida no dia. | UAN FTT |
| Semana letiva | 1 – n | Contagem de semanas até o fim do semestre. | Calendário FTT |
| Dia letivo | 1 – n | Contagem de dia até o fim do semestre | Calendário FTT |
| Temperatura mínima | Valor numérico | Temperatura mínima apresentada no dia vigente. | Instituto Nacional de Meteorologia |
| Precipitação binária | Variável categórica | 1 se choveu e 0 caso não tenha chovido no dia | Instituto Nacional de Meteorologia |
| Tipo preparo | Variável categórica | Modo de preparo, varia entre cozido, assado, frito e grelhado. | UAN FTT |
| Tipo refeição | Variável Categórica | Varia entre carne, frango, peixe e porco. | UAN FTT |

Tabela - Tabela de variáveis

As variáveis **Dia letivo, Semana letiva, Dia do Letivo do Bimestre** e **Dia da Semana** foram utilizadas para que o modelo pudesse estabelecer uma relação de sazonalidade na frequência das pessoas no refeitório. Vale ressaltar também que nos dados de base que são usados nos testes e treinamento do código, estão sendo considerados apenas os dias letivos que foram apresentados como tal no calendário letivo fornecido pela própria faculdade.

Os dados de **Precipitação**, **Temperatura Mínima** e **Precipitação Binária** foram coletados e adicionados à base visando estudar se fatores climáticos iriam ter impacto na frequência do refeitório. Visto que instituição fica localizada em uma região central do Grande ABC, muitas vias de acesso que os alunos utilizam para se deslocar até a Faculdade podem ficar alagadas e obstruídas impedindo ou atrasando a chegada dos alunos até a faculdade.

**Guarnição, Tipo preparo e Tipo refeição** são referentes as refeições que são servidas no refeitório e seguem mais ou menos um padrão: É servido arroz, feijão, um ou dois tipos de guarnição, salada, suco ou sobremesa. Por conta da grande variedade de tipos de pratos que são servidos, foi decidido agrupá-los para que ficasse mais fácil identificar um padrão. Os alunos e colaboradores são livres para decidir se vão jantar no refeitório ou não, então pode ser que determinados tipos de pratos influenciem mais ou menos a frequência das pessoas em determinado dia.

O número de refeições servidas representa o número de pessoas que frequentaram o refeitório e é representado pela variável **Refeições na Janta**. Esse dado é registrado todos os dias e trata-se da variável dependente, ou seja, o dado que vai nortear o treinamento do algoritmo.

Para aplicação deste algoritmo foi escolhido a linguagem de programação Python com a biblioteca *Sklearn*, que é uma biblioteca de aprendizado de máquina de código aberto que suporta aprendizado supervisionado e não supervisionado. Essa biblioteca também fornece várias ferramentas para ajuste de modelo, pré-processamento de dados, seleção e avaliação de modelo e outras utilidades (PEDREGOSA, 2011).

Para avaliação da qualidade do relacionamento dos dados coletados foram utilizados dois métodos estatísticos:

* Correlação de Pearson: Quantifica a força de associação linear entre duas variáveis, descrevendo o quanto uma linha reta se ajusta aos dados (SHIKAMURA, 2005).
* Correlação de Spearman: Quantifica o grau de associação de duas variáveis não lineares (PONTES, 2010).

 Ambas são valores medidos entre -1 e 1, o primeiro valor um relacionamento inversamente proporcional enquanto o segundo um relacionamento diretamente proporcional (SHIKAMURA, 2005).

**3.1 Treinamento e Otimização do Algoritmo**

Para treinar o algoritmo, a base foi dividida em duas categorias: Dados para treinamento e dados para teste. A base de testes é criada através da seleção aleatória de alguns registros da base de dados, e possui 20% dos dados enquanto o treinamento do algoritmo é feito com os outros 80%. Essa opção de se utilizar este alto percentual da base para treinamento teve relação com a quantidade de amostras coletadas*.* Foram feitos vários sorteios aleatórios para tentar chegar a uma homogeneidade da base.

A figura 4 demonstra a incidência de registros de cada semana letiva na base utilizada para testes.

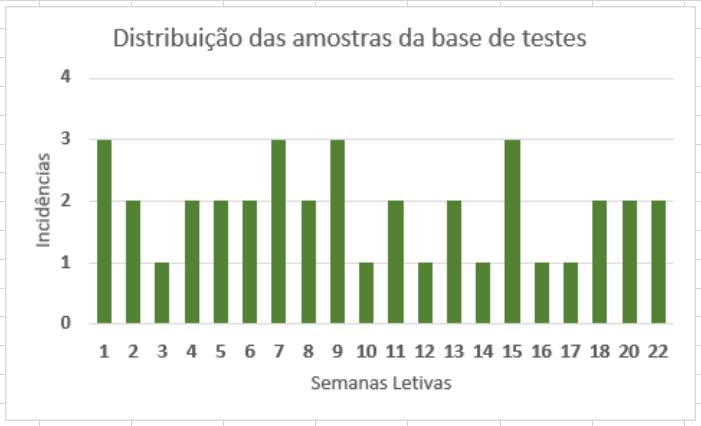


Figura 4 – Gráfico de distribuição da base de testes

Ao longo do treinamento, foi realizada a otimização do algoritmo utilizando tanto o *Grid Search*, que ocorre antes da sua execução, como o *Permutation Feature Importance*, que ocorre após a sua conclusão.

Tendo o modelo treinado, os dados da base de teste foram aplicados e os resultados foram coletados para reportar dentro da pesquisa.

1. **Resultados obtidos**
   1. **Estudo de correlação**

O estudo de correlação da variável “refeições na janta” com as demais não apresentou nenhuma correlação forte em ambos os métodos analisados, conforme é possível ver na Tabela 2. Esse fato não indica que não exista correlação, somente que não existe relação linear ou monótona entre as variáveis.

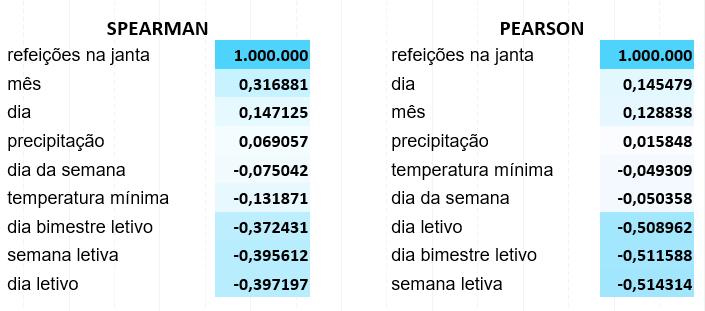


Tabela - Resultado de estudo de correlação

Na figura 5 observam-se uma série de gráficos que foram utilizados para a inspeção e verificação visual da relação das variáveis, e nela é possível perceber que alguns gráficos são muito dispersos, não apresentando nenhuma forma ou padrão definido.

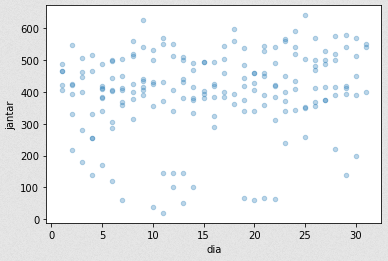
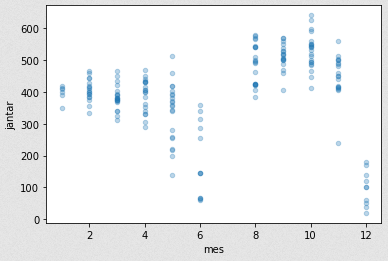
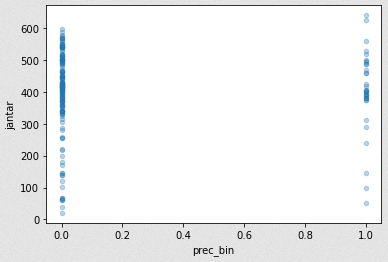
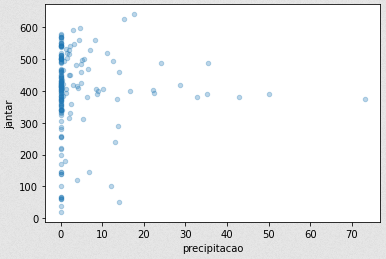
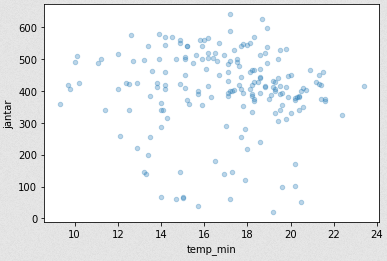
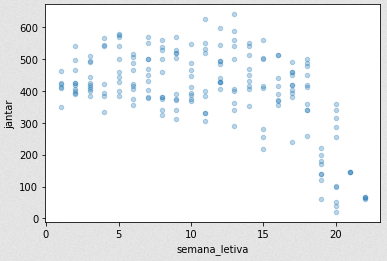
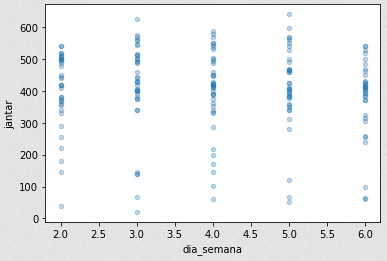
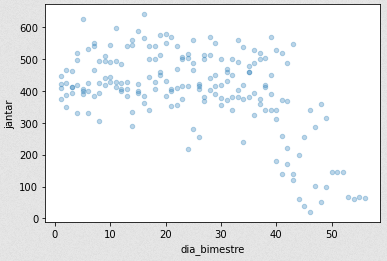
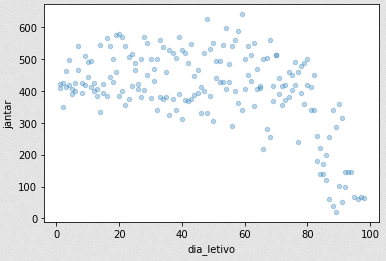


Figura 5 - Gráfico de inspeção visual das variáveis

No entanto, pode-se observar que as variáveis referentes ao calendário letivo (dia\_letivo, dia\_bimestre, semana letiva) indicam uma queda significativa na quantidade de refeições servidas no final do gráfico. Também se pode observar, através da variável Mês, um aumento da quantidade de refeições servidas no segundo semestre do ano observado, 2018.

* 1. ***Permutation feature importance***

Os resultados da aplicação do método de *Permutation feature importance* indicaram que uma boa parte das variáveis não era bem aproveitada pelo modelo para estimar as refeições. Umas pelo baixo valor calculado, e outras pelo fato de o valor estar muito ligado ou próximo ao desvio padrão calculado, como se pode ver na tabela 3:



Tabela - Resultados do Permutation Feature Importance

Assim sendo, foi decidido retirar do modelo as que não apresentaram valores acima de 5% de relevância no modelo. Portanto, as variáveis utilizadas para elaboração do modelo foram: dia leito bimestre, semana letiva, dia letivo e temperatura mínima.

* 1. **Dados previstos**

Utilizando as variáveis selecionadas, foi realizada a criação do algoritmo, que se encontra disponível em [<https://github.com/SrMartis/TCC1/blob/master/tcc_clean.ipynb>]. Por questões de sigilo, a base utilizada como referência não pode ser divulgada neste caso.

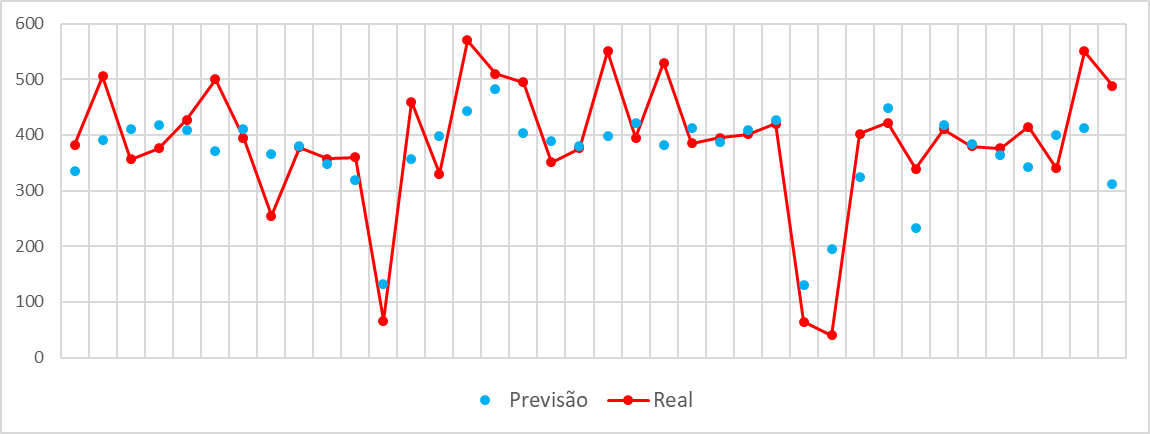
Os dados que foram previstos pelo algoritmo permitiram a geração do gráfico 6, onde a linha em vermelho liga os pontos que representam a quantidade de refeições servidas e os pontos em azul são a estimativa de refeições que deveriam ser preparadas de acordo com o modelo treinado.

Figura 6 - Comparação dos dados previstos com os reais

Ao se analisar o gráfico é possível verificar que o modelo de aprendizado de máquina é responsivo aos dados de base inseridos e tenta gerar resultados próximos aos reais, no entanto em alguns pontos acabam ficando muito distantes.

* 1. **Coeficiente de determinação**

O algoritmo SVR e algoritmos de regressão em geral são medidos através de uma métrica de *score*, no caso o coeficiente de determinação, também conhecido como R², que varia de 1 a 0. Essa métrica expressa a quantidade de variância do modelo, e quanto maior o valor de R², melhor o modelo se ajusta aos dados.

No melhor teste executado, o algoritmo gerou um coeficiente de determinação de 0.54. Este número indica que o modelo aplicado e ajustado, explica 54% da variância do resultado a partir das variáveis estudadas.

* 1. **Análise estatística**

Ao se pegar os dados que o modelo calculou e compará-los aos dados reais, é possível aplicar alguns conceitos de estatística para analisar e entender melhor estes resultados, na tabela 4 temos alguns destes:

|  |  |
| --- | --- |
| **Descrição** | **Refeições** |
| Erro absoluto médio | 61,9 |
| Erro absoluto mediano | 41,3 |
| Erro máximo | 189,6 |
| Erro absoluto médio para mais | 41,9 |
| Erro absoluto médio para menos | 81,8 |
| Erro absoluto mediano para mais | 32,8 |
| Erro absoluto mediano para menos | 74,3 |
| Erro máximo para mais | 189,6 |
| Erro máximo para menos | 173,0 |
| Desvio padrão para mais | 44,3 |
| Desvio padrão para menos | 49,5 |
| Desvio padrão de erro | 51,1 |

Tabela - Resultados estatísticos

Para analisar esses dados, é necessário ter conhecimento que o valor médio de refeições servidas no ano analisado foi 403. Assim sendo, o erro absoluto médio, uma das principais métricas que indicam o quanto o algoritmo erra, representa cerca de 15% quando comparado a média de refeições servidas. Pode não parecer muito, mas deve-se lembrar que neste assunto, se trata de desperdício em alguns dias, ou falta de comida em outros.

Separando os casos de desperdício dos casos de falta, nota-se que o algoritmo parece não exagerar na hora de fazer as predições, já que o erro médio absoluto para mais, é muito menor do que o erro para menos, o que é um aspecto bom do modelo, pois pode-se criar uma alternativa para uma possível falta, mas a comida que sobra nesses tipos de refeitórios não pode ser reaproveitada e tem que ser descartada. No entanto, o modelo faz predições muito baixas em várias oportunidades, gerando um erro absoluto mediano para menos igual a 74 refeições e por tratar-se de uma mediana, o número esconde valores muito maiores que esse.

Também chama atenção os erros máximos que foram registrados, tanto para mais, quanto para menos. Os valores foram extremamente altos, e qualquer pessoa que olhar para estes números vai achar preocupante, e com razão. Porém, principalmente o desvio padrão, indica que estes são pontos fora da curva e para notar isso basta olhar para qualquer outro valor de média na tabela, mas também é possível constatar isso, no fato do erro absoluto mediano geral ser cerca de 20 refeições a menos que erro absoluto médio.

1. **Considerações Finais**

Apesar de o objetivo da pesquisa estar direcionado ao estudo do SVR como ferramenta de apoio na decisão da quantidade de refeições servidas em um refeitório e tal objetivo ter sido atingido, diante dos resultados obtidos, não é conclusivo dizer que o modelo possa substituir a avaliação empírica do gestor.

Deve-se levar em consideração que o resultado foi fortemente influenciado pelas características específicas do refeitório analisado e pela quantidade de dados fornecidos, que caracterizam limitações da pesquisa. Mesmo diante deste cenário, o modelo foi capaz de acompanhar de maneira razoável a curva real de demanda. Isso pode ser um forte indício que os resultados possam ser melhorados em estudos futuros.

É possível que o algoritmo treinado com uma base de dados maior, com mais variáveis, tenha resultados mais assertivos, o que se sugere como um trabalho futuro. Para tanto, seria necessário a obtenção de novos dados para compor a base utilizada no estudo e reavaliar se isso torna possível a utilização do modelo como ferramenta de apoio.

1. **Referencial Bibliográfico**

ABREU, E.S.; SIMONY, R.F.; DIAS, D.H.S.; RIBEIRO, F.R.O. Avaliação do desperdício alimentar na produção e distribuição de refeições de um hospital de São Paulo. Simbologias, v.5, n.7, p.42-50. 2012.

AGGARWAL, Charu C. (Ed.). Data classification: algorithms and applications. CRC press, 2014.

AWAD, Mariette; KHANNA, Rahul. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. [S. l.]: Apress, 2015. 268 p. ISBN 9781430259909. E-book.

AWAD, Mariette; KHANNA, Rahul. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. [S. l.]: Apress, 2015. p 4. ISBN 9781430259909. E-book.

CHARNIAK, E. e McDERMOTT, D. (1985). Intradaylion to Artificial Intelligence. Addison-Wesley.

CHATTERJEE, Samprit; HADI, Ali S. Regression analysis by example. John Wiley & Sons, 2015.

COLOG, Boas Práticas na Manipulação de Alimentos, Gestão das Sobras – Como reduzir às Sobras, 2018;

CFN - Conselho Federal de Nutricionistas. RESOLUÇÃO CFN Nº 600, DE 25 DE FEVEREIRO DE 2018. Diário Oficial da União. 2018. Disponível em: https://www.cfn.org.br/wp-content/uploads/resolucoes/Res\_600\_2018.htm

FAO - Food Agriculture Organization of the United Nations. Food Loss and Food Waste. 2018?. Disponível em: <http://www.fao.org/food-loss-and-food-waste/en/>

FEIGENBAUM, A. E. Expert Systems, principles and practice in the Encyclopedia of computer Science and Engineering, 1992.

FERNANDES, Anita Maria da Rocha - Inteligência Artificial noções gerais – 2°, Florianópolis: VisualBooks, 2005.

HAUGELAND, John. Artificial Intelligence: The Very Idea. Massachusetts: The MIT Press, 1985.

HOSCH, L. WILLIAM. Machine learning. Enciclopedia Britannica. 2020. Disponível em:  
https://www.britannica.com/technology/machine-learning, Acesso em: 16/06/2020.

JUNIOR, E. A.; TEIXEIRA, R. P. A. Manual de procedimentos para utilização de sobras alimentares. Modelo nutrição: módulo programação. Rio de Janeiro: SESC, 2010.

PAULA, Ruben. Máquinas de suporte vetorial como instrumento de priorização de investimentos aplicado ao Programa de Investimento e Logística – Aeroportos. 268 p. 2016.

PEDREGOSA et al. Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830, 2011.

PEIXOTO, M.;PINTO, H. S. Desperdício de Alimentos: questões socioambientais, econômicas e regulatórias. Brasília: Núcleo de Estudos e Pesquisas/CONLEG/ Senado (Boletim Legislativo nº 41, de 2016). Disponível em: www.senado.leg.br/estudos. Acesso em 11/05/2020.

PEREIRA, D. X. R. Going zero waste in canteens: Exploring food demand using data analytics. Faculdade de Engenharia da Universidade do Porto, 18 de julho de 2018.

PONTES, A.C.F (2010). Ensino da correlação de postos no ensino Médio.   
19º SINAPE - Simpósio Nacional de Probabilidade e Estatística. Associação Brasileira de Estatística.  
26 a 30 de julho de 2010. São Pedro-SP.

SENAI, INSTITUTO DE INOVAÇÃO SENAI. APRENDIZAGEM SUPERVISIONADA OU NÃO SUPERVISIONADA? Disponível em: https://isitics.com/2018/08/28/aprendizagem-supervisionada-ou-nao-supervisionada/, Acesso em: 16/06/2020.

SCOTTON, V. et al. Desperdício de Alimentos em Unidades de Alimentação e Nutrição: a contribuição do resto-ingestão e da sobra. Revista Higiene Alimentar, v. 24, n. 186/187, p. 19-24, 2010.

SHIMAKURA, SILVIA EMIKO. UFPR - Laboratório de Estatística e Geoinformação.  
2005. Disponível em: http://www.leg.ufpr.br/~silvia/CE701/node79.html.  
Acesso em: 18/06/2020.

TEIXEIRA, S. M. F. G. et al. Administração aplicada às unidades de alimentação e nutrição. São Paulo: Atheneu, 2007.

WINSTON, P. H. (19921, Artificial Intelligence (Third -edition) Addison-Wesley.