**Comparação Entre o uso de Redes Neurais e Regressões Múltiplas para Custo de Matéria Prima**

David Santos Baião;Maurício Acconcia Dias

1 Cummins Filtros Brasil. Universidade de São Paulo. ESALQ. Piracicaba, Estado de São Paulo - Brasil. Rua Augustinho Soares, 78 – Horto 3; 18074-795 Sorocaba, SP, Brasil

2 Centro Universitário Herminio Ometto de Araras. Avenida Doutor Maximiliano Baruto Jardim Universitário; 13607-339 - Araras, SP – Brasil

\*autor correspondente: david\_baiao@yahoo.com.br

**Comparação Entre o uso de Redes Neurais e Regressões Múltiplas para Custo de Matéria Prima**

**Resumo**

A previsão do custo de matéria prima é um processo crucial para que as empresas sejam capazes de planejar preços de vendas e para o planejamento financeiro de curto prazo. Este trabalho destina-se a analisar e comparar os métodos regressão múltipla utilizando do software Minitab® e redes neurais utilizando o software R® na previsão de custo de matéria prima, utilizando índices de mercado como o preço do aço, dólar, e índice de inflação.

**Palavras-chave: Redes Neurais, Regressões Múltiplas, Custo de Matéria Prima, Minitab, R, Indices de Mercado**

**Comparison Between the Use of Neural Networks and Multiple Regressions for Raw Material Cost**

**Abstract**

The forecast of the raw material cost it is a crucial process for companies to be able to create the sales prices plan and short-term financial planning. This work intends to analyze and compare the methods of multiple regression using the Minitab® software and neural networks using the R® software in the forecast of raw material cost, using market indices such as the price of steel, foreign exchange rate, and inflation Index.

**Neural Networks, Multiple Regressions, Cost of Raw Materials, Minitab, R, Market Indices**

**Introdução**

A previsão de custo de matéria prima é um processo crucial para qualquer empresa desenvolver os planos de negócios a curto prazo, pois esse é um dos fatores que impactam diretamente as margens de lucro.

É amplamente utilizado no mercado o modelo de custo direto e indireto no qual o custo direto que engloba os custos de aquisição de matéria prima e custos fabris e administrativos que variam diretamente com a quantidade produzida e os custos indiretos que não variam diretamente com o volume produzido.

Cada empresa tem uma variação dessas nomenclaturas bem como suas próprias regras de categorização de cada desembolso no cálculo de custeio, no entanto o custo de matéria prima é em regra geral o mais simples de ser classificado, pois nele contém todos os componentes (parafusos, componentes eletrônicos etc.), materiais brutos (Chapas de metal, Plásticos etc.) e os custos de transformação, transporte e comercialização que são utilizados para a produção e compra do item.

Em uma visão geral é comum saber que cada um dos grupos que compõe o custo do produto tem uma variação distinta, pois são influenciados por diferentes fatores e alguns deles variam independentemente de como a administração da empresa é feita, como as variações de comodities no mercado que têm impacto direto no custo de matéria de seus subfornecedores, acordos sindicais que influenciam o custo de mão de obra.

Nesse trabalho será analisada a relação que a variação da comodity do Aço, bem como índices como IPCA e valor do Dólar tem em um filtro amplamente utilizado no mercado brasileiro.

Para essa análise serão utilzidas ferramentas estatísticas como a regressão múltipla no software Minitab®, e redes neurais com o uso do software R®, a fim de identificar quais os fatores têm maior impacto bem como qual a melhor metodologia tem melhor acurácia para previsão de custo de matéria prima para o item em questão.

**Material e Métodos**

Para a coleta de dados serão utilizados os dados internos de custo matéria prima do relatório de vendas da empresa que tem dados diários de vendas desde janeiro de 2013. Os custos serão filtrados para o item em questão e calculada sua média mensal para comparação com os índices quem tem a mesma base mensal de dados.

Os índices serão coletados dos índices da FVG (Fundação Getúlio Vargas) Bobinas quente de aço ao carbono código 1420793, o Bobinas frio de aço ao carbono código1420792, IPCA código 156038 e o índice do BACEN (Banco Central do Brasil) Dólar US$ Dólar dos EUA.

Para a análise dos dados serão montados em formato de tabela utilizando a média aritmética de acordo com o período mensal.

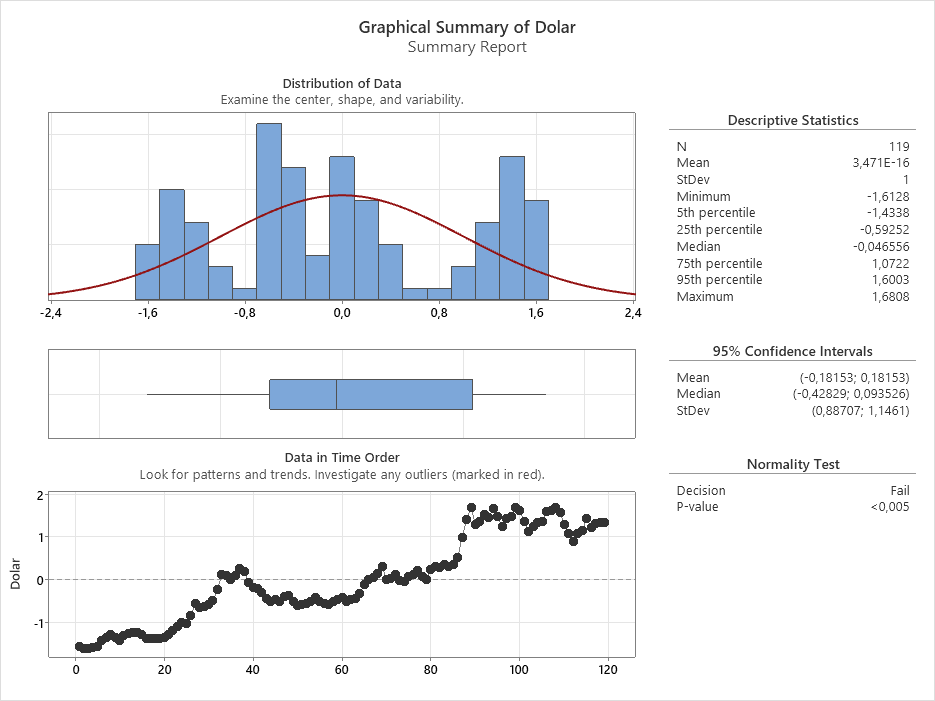
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tabela 1: Dados Históricos normalizados para de Matéria Prima e Índices de Mercado | | | | |
| Mês | Matéria Prima | IPCA | Aço | Dólar |
| 01/11/2022 | 1.76 | 1.95 | 1.81 | 1.34 |
| 01/10/2022 | 1.71 | 1.92 | 1.83 | 1.32 |
| 01/09/2022 | 1.75 | 1.87 | 1.90 | 1.31 |
| 01/08/2022 | 2.00 | 1.89 | 1.93 | 1.22 |
| 01/07/2022 | 1.74 | 1.92 | 1.95 | 1.43 |
| 01/06/2022 | 1.75 | 1.98 | 2.12 | 1.14 |
| 01/05/2022 | 1.76 | 1.92 | 2.20 | 1.06 |
| 01/04/2022 | 1.80 | 1.89 | 2.01 | 0.88 |
| 01/03/2022 | 1.89 | 1.80 | 1.84 | 1.07 |
| 01/02/2022 | 2.17 | 1.67 | 1.86 | 1.28 |
| 01/01/2022 | 2.12 | 1.59 | 1.97 | 1.57 |
| 01/12/2021 | 2.20 | 1.55 | 2.01 | 1.68 |
| 01/11/2021 | 2.27 | 1.49 | 2.09 | 1.60 |
| 01/10/2021 | 2.00 | 1.42 | 2.39 | 1.58 |
| 01/09/2021 | 1.94 | 1.32 | 2.38 | 1.35 |
| 01/08/2021 | 1.92 | 1.23 | 2.35 | 1.32 |

Para analisar esses dados e suas influências é importante que seja feita uma análise dos dados da Tabela 1, e como pode ser visto nas figuras abaixo (Figura 1: Matéria Prima, Figura 2: IPCA, Figura 3: Aço e Figura 4: Dólar), todos os itens da tabela demonstram crescimento ao longo do tempo o que demostra chances de correlação entre elas.

|  |
| --- |
| Figura 1: Matéria Prima  Gráfico  Descrição gerada automaticamente |
|  |

|  |
| --- |
| Figura 2: IPCA |
|  |
| Figura 3: Aço |
|  |

Figura 4: Dólar



**Resultados e Discussão**

## Análise de regressão utilizando Minitab®

Utilizando a correlação de Pearson somos capazes de identificar a influência que cada item tem com os demais, e com isso podemos verificar se se o IPCA, o Aço e o Dólar têm realmente correlação com a Matéria Prima e uns com os outros (Figura 5: Correlação de Pearson para os itens Matéria Prima, IPCA, Aço e Dólar).

Figura 5: Correlação de Pearson para os itens Matéria Prima, IPCA, Aço e Dólar

Gráfico, Gráfico de dispersão

Descrição gerada automaticamente

Com base na Figura 5, pode-se verificar que todos os itens tem correlação forte e positiva com a Matéria prima (IPÇA com r=0,902, Aço com r= 0,908 e Dólar com r=8,854), e podemos também verificar forte correlação entre eles o que indica que podemos ter itens dependentes uns dos outros e nesse caso pode-se retirar um deles da análise.

Para definir se manteremos todos os indicadores na medição utilizaremos a regressão dos melhores subconjuntos que é mostrado na Figura 6.

|  |
| --- |
| Figura 6: Regressão dos Melhores Subconjuntos: Matéria Prima versus IPCA; Aço; Dólar |
|  |

Conforme destacado no Figura 6, o conjunto de índices Dolar, Aço e IPCA mostram o maior conjunto de valores para o R² (R-quad, R2(aj) e R2(pred)), sendo assim todos eles serão utilizados para previsão matéria prima em uma regressão múltipla.

Ao utilizar a análise de Regressão Múltipla feita no Minitab® 20.4, temos como resultado a precisão de 92,4% variação explicada pelo conjunto de variáveis proposto conforme pode ser visto na Figura 7: Análise de Regressão: Sumário do Modelo

Figura 7: Análise de Regressão: Sumário do Modelo

Interface gráfica do usuário, Gráfico

Descrição gerada automaticamente

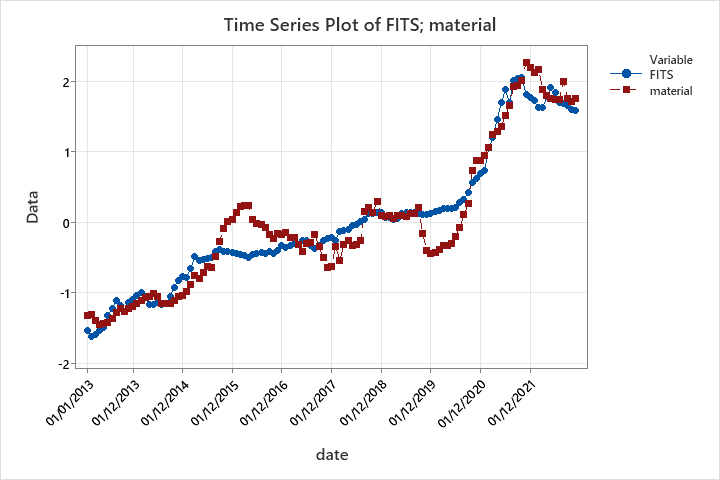
Tem-se também como resultado da análise de regressão a equação do modelo que pode ser vista na Figura 8: Análise de Regressão: Equação de Regressão.

|  |
| --- |
| Figura 8: Análise de Regressão: Equação de Regressão |
|  |

Com base nos valores da coluna “Matéria Prima” da Tabela 1: Exemplo de Dados Históricos e a Figura 8: Análise de Regressão: Equação de Regressão, podemos gerar o Figura 10: Gráfico de Matéria Prima e Equação de Regressão que mostra de maneira sobreposta os resultados da equação e os dados da coluna.

Pode ser visto também na Figura 9 que temos dois períodos que mostra maior divergência entre os valores da calculados e os valores reais, mostrando que quando há variações abruptas dos valores de matéria prima a regressão tem menor precisão nos resultados.

Figura 9: Gráfico de Matéria Prima e Equação de Regressão



## Análise por redes neurais utilizando o RStudio®

Para comparamos a técnica de Regressão Múltipla e as Redes Neurais será utilizado do software RStudio, Versão 1.4.1717.

Inicialmente o deve-se fazer o carregamento dos dados. Após isso o carregamento dos dados para análise e a preparação do banco de dados de análise.

Então torna-se necessário separar os dados em amostra e teste para que seja possível analisar a eficácia do resultado que para esse projeto será utilizado 20% como amostra de maneira aleatória ao longo do período da amostra.

Para a análise por redes neurais será utilizado o pacote Neuralnet utilizando o IPCA e Aço para fazer a construção da rede neural com 24 neurônios em 4 camadas e fazendo 100 réplicas de estudo para a construção do modelo.

Com o de rede neural “NN\_00” criado na etapa anterior pode-se calcular os resultados utilizando os bancos de dados de treino, de teste e o total para compararmos com os valores reais contidos na coluna Material.

A partir dos dados preditos anteriormente é possível fazer a análise de correlação de Pearson para compararmos com os resultados obtidos no método de regressão.

Com os comandos abaixo podemos ver os resultados da utilização de redes neurais, que segundo a correlação de Pearson o modelo com a tabela de treino tem 99,9% de correlação, já o para a tabela de teste tem 98,2% de correlação, e para a tabela total com todos os valores têm 98,9% de correlação.

Em posse dos dados obtidos através da predição do modelo da rede neural, pode-se construir um gráfico para visualizar de maneira sobreposta os dados dos valores de matéria prima e as predições (Figura 11: Gráfico de Matéria Prima e Predição do modelo de rede neural).

Figura 10: Gráfico de Matéria Prima e Predição do modelo de rede neural

**Gráfico, Gráfico de linhas

Descrição gerada automaticamente**

**Conclusão(ões) ou Considerações Finais**

Fazendo a comparação entre os métodos podemos verificar que a rede neural (R² = 0.989) teve resultado mais preciso que a regressão múltipla (R²=0,93) e que graficamente podemos identificar que essa superioridade aparece nos postos em que há variação abrupta do valor de matéria prima seja para aumentar ou para reduzir.

**Referências**

Fávero, Luiz Paulo; Belfiore, Patrícia. (2017). Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®. Rio de Janeiro: Elsevier

Wickham, H. & Grolemund, G. R for Data Science: <https://r4ds.had.co.nz/index.html>.

Barbara F. Ryan; Brian L. Joiner; Jonathan D. Crye. (2012). Minitab Handbook: Update For Release 16. California: Thomson Brooks/cole

**Apêndice ou Anexo** (opcional)

Apêndices são textos e/ou documentos que foram elaborados pelo autor e que são importantes para complementar a argumentação do trabalho. Anexos são textos ou documentos que ilustram, mas que não foram elaborados pelos autores. Apêndices deverão seguir as mesmas normas de formatação do restante do texto, inclusive para figuras e tabelas.

O TCC deverá conter no máximo 30 páginas, incluindo o(s) Apêndice(s) e/ou Anexo(s).