

Universidade Federal do Paraná
Setor de Ciências Exatas
Departamento de Estatística
MBA em *Advanced Analytics & Business Optimization*

João Eduardo Felipin

**Antecipando a Demanda de materiais de
limpeza hospitalar: Avaliação de Modelos de
Séries Temporais**

Curitiba
2025

João Eduardo Felipin

Antecipando a Demanda de materiais de limpeza hospitalar: Avaliação de Modelos de Séries Temporais

Monografia apresentada ao MBA em *Advanced Analytics & Business Optimization* da Universidade Federal do Paraná como requisito parcial para a obtenção do grau de especialista.

Orientador: Prof. Guilherme Vinicyus Batista

Curitiba
2025

Antecipando a Demanda de materiais de limpeza hospitalar: Avaliação de Modelos de Séries Temporais

João Eduardo Felipin¹, Guilherme Vinicyus Batista²

¹Departamento de Gestão da Informação, Universidade Federal do Paraná, Av. Prefeito Lothário Meissner, 632, Jardim Botânico - Curitiba/PR - 80210-070*

²Professor Convidado - UFPR, guivbatista89@gmail.com

A gestão eficiente de insumos hospitalares é fundamental para garantir a sustentabilidade operacional e a qualidade dos serviços de saúde. Em ambientes caracterizados por alta variabilidade de consumo e complexidade operacional, a previsão de demanda baseada em séries temporais torna-se uma ferramenta estratégica para apoiar a tomada de decisão. Este estudo propõe a aplicação e comparação de diferentes técnicas de previsão para estimar o consumo de materiais hospitalares, com foco em insumos de limpeza. Os dados foram extraídos do sistema corporativo de gestão de almoxarifado do hospital e submetidos a procedimentos de tratamento, incluindo correção de outliers por meio do método do intervalo interquartil e agregação temporal. Foram avaliados modelos estatísticos tradicionais (ARIMA e SARIMA), média móvel, redes neurais LSTM e uma abordagem de rolling forecasting baseada em LSTM. Os resultados indicam que o modelo LSTM com rolling forecasting apresentou o melhor desempenho, com menores erros de previsão, evidenciando maior capacidade de adaptação a mudanças recentes no padrão de consumo e maior adequação ao contexto hospitalar.

Palavras-chave: Previsão de demanda, logística hospitalar, aprendizado de máquina

Efficient hospital supply management is essential to ensure operational sustainability and maintain health-care service quality. In environments characterized by high consumption variability and operational complexity, time series demand forecasting becomes a strategic tool to support decision-making. This study proposes the application and comparison of different forecasting techniques to estimate the consumption of hospital supplies, focusing on cleaning materials. The data were extracted from the hospital's inventory management system and subjected to preprocessing procedures, including outlier correction using the interquartile range method and temporal aggregation. Traditional statistical models (ARIMA and SARIMA), moving averages, Long Short-Term Memory (LSTM) neural networks, and an LSTM-based rolling forecasting approach were evaluated. The results show that the LSTM model with rolling forecasting achieved the best predictive performance, presenting lower forecast errors and greater adaptability to recent changes in consumption patterns, demonstrating its suitability for dynamic hospital environments.

Keywords: Demand forecast, hospital logistics, machine learning

1. Introdução

A gestão eficiente de insumos hospitalares é um fator crítico para a sustentabilidade operacional e a qualidade do atendimento em instituições de saúde. Em um contexto marcado por alta complexidade, variabilidade de consumo e restrições orçamentárias, torna-se essencial adotar métodos quantitativos capazes de apoiar o planejamento e a tomada de decisão base-

ada em dados históricos. Nesse cenário, a previsão de demanda por meio de séries temporais surge como uma ferramenta estratégica, permitindo antecipar variações no consumo de materiais e reduzir riscos associados à falta ou ao excesso de estoque. Este projeto propõe a aplicação e comparação de diferentes técnicas de previsão de séries temporais em ambiente hospitalar, analisando seus desempenhos por meio de métricas quantitativas e avaliando sua adequação frente às características específicas desse tipo de dado.

*joaoed.felipin@gmail.com

1.1. Previsão no ambiente hospitalar

Dentro do ambiente hospitalar o conceito de previsão vem sendo aplicado para a antecipação e estudos de doenças, sendo essas técnicas utilizadas em conjunto de variáveis externas, como demonstra Martins et al. (2025). Quando observamos no âmbito operacional, podemos aplicar em diversas frentes como a limpeza de áreas, ocupação de leitos, manutenção. Isso reforça que quando um hospital tem a capacidade de flexibilizar e adaptar os seus processos operacionais, o mesmo sofre um efeito positivo (Naser et. al., 2023).

Um exemplo prático é o estudo realizado no Bushehr Heart Hospital, no Irã, em que a simulação de eventos foi utilizada para diagnosticar e otimizar o fluxo de pacientes. Essa abordagem permitiu identificar causas de filas e sobrecarga de trabalho, propondo soluções baseadas em modelagem e otimização de processos, o que resultou em melhoria no tempo de atendimento e no bem-estar das equipes (Tello et al., 2022).

Sendo assim, podemos notar que o uso de ferramentas de previsão no contexto operacional hospitalar não apenas melhora a alocação de recursos, mas também contribui para reduzir custos, consolidando uma cultura de tomada de decisão baseada em dados.

1.2. LSTM (Long Short-Term Memory)

As Redes Neurais Artificiais (RNAs) constituem uma das principais classes de modelos de aprendizado de máquina, inspiradas no funcionamento do cérebro humano. A ideia central dessas redes é a capacidade de aprender padrões por meio do ajuste dos pesos das conexões entre unidades chamadas neurônios artificiais (HAYKIN, 1999). Esses modelos são capazes de aproximar funções complexas, identificar relações não lineares e generalizar comportamentos a partir de exemplos observados.

Dentro das arquiteturas clássicas de redes neurais, as Redes Neurais Recorrentes (RNNs) representaram um avanço importante ao introduzir um mecanismo de memória interna. Diferentemente das redes *feed-forward*, que processam as entradas de forma independente, as RNNs mantêm estados ao longo do tempo, permitindo a modelagem de dependências sequenciais (RUMELHART; HINTON; WILLIAMS, 1986). Essa característica tornou as RNNs especialmente adequadas para tarefas como:

- previsão de séries temporais;
- processamento de linguagem natural;
- reconhecimento de padrões sequenciais;

- modelagem de dependências temporais em dados clínicos.

Apesar dessas vantagens, as RNNs tradicionais apresentam limitações significativas, principalmente associadas aos problemas de *vanishing gradient* e *exploding gradient*, que dificultam o aprendizado de relações de longo prazo em sequências extensas.

Foi nesse contexto que Hochreiter e Schmidhuber (1997) propuseram as redes *Long Short-Term Memory* (LSTM), uma variação das RNNs desenvolvida para lidar de forma mais eficiente com dependências temporais de longo alcance. O principal diferencial das LSTM está na introdução de mecanismos de controle do fluxo de informação, conhecidos como *gates* (portas):

- *input gate* — controla quais informações devem ser adicionadas à memória;
- *forget gate* — define quais informações devem ser descartadas;
- *output gate* — regula quais informações serão utilizadas como saída.

Esses mecanismos permitem que a rede preserve informações relevantes por períodos prolongados, evitando o esquecimento prematuro e tornando o processo de aprendizado mais estável. Em razão dessas características, as LSTM tornaram-se uma das arquiteturas mais populares para aplicações envolvendo séries temporais, incluindo domínios como saúde, previsão de demanda hospitalar e análises operacionais.

A evolução das redes neurais, no entanto, não se limitou às LSTM. Arquiteturas mais recentes, como as *Gated Recurrent Units* (GRU) (CHO et al., 2014) e modelos baseados em mecanismos de atenção, como os *Transformers* (VASWANI et al., 2017), também foram desenvolvidas com o objetivo de capturar dependências de longo prazo. Ainda assim, as LSTM permanecem amplamente utilizadas devido à sua estabilidade, interpretabilidade relativa e bom desempenho em séries temporais estruturadas, como aquelas encontradas em ambientes hospitalares.

1.3. ARIMA e SARIMA

Quando falamos de dados distribuídos no tempo, não podemos fugir de explicar as séries temporais. Esses modelos constituem a base das abordagens estatísticas tradicionais para previsão em séries temporais, utilizando padrões históricos para estimar valores futuros. Antes da popularização de métodos de aprendizado

de máquina, modelos como ARIMA e SARIMA eram amplamente adotados como padrão em estudos de previsão (BOX; JENKINS; REINSEL, 2008).

1.3.1. ARIMA

O método ARIMA é um modelo tradicional da estatística que significa “autorregressivo integrado de médias móveis”. Segundo o livro de BOX; Jenkins e Reinsel O ARIMA é composto por três componentes:

$$\text{ARIMA}(p, d, q) = \text{AR}(p) + \text{I}(d) + \text{MA}(q)$$

AR (Componente Autorregressivo):

Consiste em utilizar os valores prévios que aparecem na série temporal para prever os valores futuros, partindo do pressuposto de que existe um padrão nos dados passados que ajuda a prever o futuro.

$$Y_t = c + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \cdots + \varphi_p \varepsilon_{t-p} + \varepsilon_t$$

Onde:

φ_i são coeficientes autorregressivos;

ε_t é um erro aleatório (ruído branco), com média zero e variância constante.

I (Componente Integrado):

Esse componente serve para deixar a série mais estável com o passar do tempo, consistindo em remover tendências, garantindo média e variância constante ao longo do tempo.

$$Y'_t = Y_t - Y_{t-1}$$

MA (Componente Média Móvel):

Essa parte utiliza os erros do passado para prever o próximo valor da série temporal, sendo uma combinação linear de erros.

$$Y_t = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Onde:

Y_t representa o valor observado da série temporal no instante t ;

c é uma constante que representa o termo médio da série;

θ_j são os coeficientes de média móvel;

ε_t é um erro aleatório (ruído branco), com média zero e variância constante.

1.3.2. SARIMA

Segundo o livro de Box, Jenkins e Reinsel, o SARIMA é composto pela combinação do método ARIMA citado acima com componentes sazonais.

$$\text{SARIMA}(p, d, q)(P, D, Q)_s$$

P (Componente Sazonal Autorregressivo):

Assim como no AR(p), o componente sazonal autorregressivo supõe que o valor atual depende de valores anteriores, porém defasados de um ciclo completo.

$$Y_t = \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \cdots + \Phi_P Y_{t-Ps}$$

D (Componente de Diferenciação Sazonal):

Remove padrões repetitivos, deixando a série estacionária em termos sazonais, o que é necessário para aplicar as propriedades estatísticas do modelo.

$$Y'_t = Y_t - Y_{t-s}$$

Q (Componente de Média Móvel Sazonal):

Assim como a média móvel tradicional, o componente MA sazonal utiliza erros passados, mas também defasados por múltiplos do período sazonal:

$$Y_t = \Theta_1 \varepsilon_{t-s} + \Theta_2 \varepsilon_{t-2s} + \cdots + \Theta_Q \varepsilon_{t-Qs} + \varepsilon_t$$

1.4. Médias Móveis (MA)

Segundo Shumway e Stoffer (2017), a média móvel é uma técnica fundamental de suavização usada para reduzir a variabilidade de curto prazo e evidenciar padrões estruturais de uma série temporal. Ela funciona aplicando uma janela deslizante que calcula a média dos valores mais recentes, permitindo identificar tendências e ciclos sem a interferência de ruídos aleatórios.

1.5. Rolling Forecasting

O rolling forecasting é uma forma de gerar previsões de maneira contínua, sempre atualizando o modelo conforme novos dados chegam. Essa abordagem é especialmente útil quando a série temporal muda com frequência, seja por variações naturais, instabilidade dos dados ou alterações no comportamento ao longo do tempo.

De acordo com Inoue, Jin e Rossi (2017), a grande vantagem desse método é que ele considera que os parâmetros estatísticos não são fixos — eles podem evoluir. Por isso, o desempenho do modelo depende muito do tamanho da janela escolhida: ela precisa ser grande o suficiente para capturar as relações importantes da série, mas também deve evitar incluir dados antigos que já não têm relação com o cenário atual.

Na prática, essa estratégia costuma se sair melhor do

que modelos que usam todo o histórico, principalmente quando a série passa por rupturas, mudanças de tendência ou apresenta outliers. A cada nova previsão, o modelo é reestimado usando apenas os dados mais recentes, o que permite que ele acompanhe rapidamente alterações na demanda, sazonalidade ou qualquer evento externo que impacte o comportamento da série.

2. Desenvolvimento

2.1. Conjunto de dados

Os dados utilizados neste estudo são provenientes do software corporativo responsável pela gestão dos almoxarifados do hospital. Esse sistema centraliza todas as movimentações relacionadas ao consumo de materiais, permitindo a extração de informações estruturadas em duas tabelas principais, denominadas neste trabalho como "tabela consumo" e "dimensão produtos". A "tabela consumo" reúne o histórico de utilização dos materiais, contendo registros de todas as saídas do estoque para as operações internas de limpeza. Entre suas variáveis, destacam-se: data do consumo, código do item e quantidade utilizada, o que possibilita analisar o comportamento temporal e operacional do uso dos materiais no hospital.

Por outro lado, a "dimensão produtos" contém as informações cadastrais dos itens presentes no almoxarifado. Essa tabela funciona como um repositório de atributos descritivos, incluindo código do produto, nome, categoria, custo unitário, unidade de medida, e outras características relevantes para a identificação e segmentação dos itens. A junção entre as duas tabelas permite contextualizar cada consumo dentro de seu grupo de produtos, fornecendo um panorama completo para análises e aplicação de técnicas de previsão. A organização nesse formato possibilita a construção de séries temporais por item, favorecendo a aplicação de modelos estatísticos e de aprendizado de máquina para prever o uso futuro de materiais hospitalares.

2.2. Definição do problema

O objetivo central deste estudo é prever a quantidade de unidades de materiais hospitalares que serão utilizadas em um período futuro previamente definido. Essa previsão é fundamental para apoiar decisões de planejamento e gestão de estoques, permitindo reduzir rupturas, otimizar compras e manter a disponibilidade contínua dos itens essenciais às atividades hospitala-

res.

A variável alvo (target) corresponde ao consumo futuro de cada item, medido em unidades, e sua previsão é formulada como um problema de série temporal. A partir do comportamento histórico de consumo, extraído da "tabela consumo", busca-se identificar padrões, sazonalidades, oscilações e tendências que possam orientar a estimativa da demanda nos períodos subsequentes.

Considerando a natureza dinâmica do ambiente hospitalar, onde fatores como sazonalidade, mudanças operacionais e flutuação no volume de atendimentos podem influenciar o consumo, a previsão torna-se um elemento essencial para a gestão eficiente do almoxarifado. Assim, o presente trabalho busca avaliar modelos capazes de capturar essas variações e fornecer previsões com assertividade, contribuindo para uma tomada de decisão mais assertiva e baseada em dados.

2.3. Implementação dos modelos

2.3.1. Preparação dos dados

Com a série diária consolidada, foi aplicado um procedimento para identificação e correção de valores anômalos (outliers). Utilizou-se o método do intervalo interquartil (IQR), no qual inicialmente são calculados os quartis Q1 e Q3. A partir deles, determina-se o IQR (Q3-Q1) e, com base nesse intervalo, calculam-se os limites inferior e superior que definem o intervalo aceitável da série. Observações que ultrapassam esses limites são consideradas outliers.

Em vez de eliminar tais registros, o que poderia comprometer a continuidade da série temporal, optou-se por substituí-los pela mediana de uma janela móvel de três dias. Esse procedimento preserva a consistência temporal do conjunto de dados, ao mesmo tempo que reduz oscilações artificiais que poderiam prejudicar o desempenho do modelo de previsão.

Como o objetivo do estudo é realizar previsões em periodicidade mensal, a série diária tratada foi posteriormente agregada por mês, por meio da soma do consumo total de cada período. Essa transformação resulta em uma série temporal mensal mais estável e apropriada para a aplicação de modelos. Para padronização e melhor legibilidade, a data representativa de cada período foi convertida para o primeiro dia do respectivo mês.

2.3.2. LSTM

Dado que modelos de redes neurais apresentam melhor desempenho quando operam em escalas numéricas uniformes, aplicou-se a técnica de normalização *Min-Max*, transformando os valores da série para o intervalo entre 0 e 1. Esse processo evita que diferenças de magnitude influenciem indevidamente o processo de aprendizagem.

Como redes *LSTM* dependem de janelas temporais para capturar dependências sequenciais, a série mensal foi convertida em sequências, adotando uma janela de 10 meses. Assim, para prever o valor do mês seguinte, o modelo recebe como entrada as dez observações mensais anteriores. Essa estrutura transforma a série em uma matriz tridimensional no formato exigido pela *LSTM*:

(samples, timesteps, features)

na qual *samples* corresponde ao número total de janelas temporais formadas a partir da série original; *timesteps* representa a quantidade de observações consecutivas utilizadas como entrada do modelo em cada janela temporal, sendo neste estudo definido como igual a 10; e *features* indica o número de variáveis consideradas em cada instante de tempo, assumindo valor igual a 1, correspondente ao consumo mensal.

O modelo construído apresenta arquitetura simples e eficiente para séries univariadas. Ele é composto por uma camada *LSTM* com 50 neurônios, utilizando função de ativação *ReLU* e formato de entrada (10, 1), seguida por uma camada *Dense* com um único neurônio, responsável pela geração da previsão do próximo valor da série.

A rede foi treinada utilizando o otimizador *Adam* e a função de perda *Mean Squared Error (MSE)*, escolhas adequadas para problemas de regressão e séries temporais contínuas.

2.3.3. ARIMA/SARIMA

Além da aplicação de modelos baseados em redes neurais, este estudo avaliou o desempenho de duas abordagens estatísticas amplamente utilizadas em séries temporais: os modelos *ARIMA* e *SARIMA*. Ambos foram ajustados de forma automatizada utilizando o pacote *pmdarima*, que identifica as melhores configurações de parâmetros a partir dos dados históricos.

A série originalmente mensal foi convertida para frequência semanal, gerando um conjunto de observa-

ções mais granular, adequado para modelos que capturam periodicidade curta e padrões sazonais. Em seguida, a série foi dividida em duas partes:

- 80% para treinamento
- 20% para teste

Essa divisão permite avaliar o desempenho dos modelos em valores não utilizados durante o ajuste, garantindo uma validação mais realista.

A modelagem *ARIMA* foi realizada por meio da função *auto_arima*, que busca automaticamente os melhores valores para os parâmetros:

- *p* — ordem autorregressiva
- *d* — grau de diferenciação
- *q* — ordem da média móvel

Essa seleção é feita com base em critérios estatísticos, como o *AIC (Akaike Information Criterion)*. Após a seleção do modelo ótimo, a série de treinamento foi ajustada e foram geradas previsões para o conjunto de teste.

O modelo *ARIMA* retorna previsões univariadas, sem considerar efeitos sazonais explícitos, servindo como referência para avaliar a necessidade de modelagem mais complexa.

Para capturar possíveis padrões sazonais semanais (como flutuações recorrentes durante o ano), foi ajustado também um modelo *SARIMA*, que estende o *ARIMA* ao incluir componentes sazonais. Nesse caso, utilizou-se:

- Sazonalidade semanal: *s* = 52
- Busca automática de parâmetros não sazonais (*p, d, q*) e sazonais (*P, D, Q*)

O *auto_arima* identifica tanto a ordem não sazonal quanto a estrutura sazonal que melhor explica o comportamento da série ao longo das semanas. Após o ajuste, o modelo foi usado para prever os mesmos pontos da etapa de teste, permitindo uma comparação direta com o *ARIMA*.

2.3.4. Médias Móveis

Além dos modelos anteriores, foi avaliado o desempenho de uma abordagem estatística simples e amplamente utilizada em séries temporais: a média móvel. Esse método consiste em suavizar a série por meio da média das últimas observações dentro de uma janela fixa, reduzindo flutuações de curto prazo e destacando

tendências gerais ao longo do tempo.

Para este estudo, foi utilizada uma janela de 5 períodos mensais, o que significa que a previsão gerada para cada mês corresponde à média dos cinco consumos anteriores. A escolha dessa janela busca equilibrar suavização e responsividade: janelas maiores tendem a produzir séries mais estáveis, enquanto janelas menores se ajustam mais rapidamente a variações recentes. Após o cálculo da média móvel, os valores foram alinhados temporalmente de forma que apenas períodos com histórico suficiente fossem considerados, garantindo a correta comparação entre valores reais e estimados.

2.3.5. Aplicação do Rolling Forecasting

Para enfrentar o dilema da alta variabilidade dos dados, foi adotado o modelo de *rolling forecasting* com *LSTM*, que permite atualizar o modelo a cada nova observação, simulando uma situação mais próxima do uso real em ambiente operacional.

Como redes *LSTM* dependem de sequências de entrada, a primeira etapa consistiu em transformar a série semanal em janelas deslizantes. Para isso, foi definida uma janela de 3 semanas, de modo que:

- as três observações anteriores constituem a entrada (X);
- a quarta observação corresponde ao valor alvo (y).

Essa estratégia cria pares de treino que capturam dependências curtas da série. Formalmente, a operação produz tensores no formato:

```
(samples, timesteps=3, features=1)
```

garantindo compatibilidade com a arquitetura da *LSTM*.

Para avaliar o modelo de forma realista, a série semanal foi dividida em:

- **Treino:** todas as semanas, exceto as últimas quatro;
- **Teste:** as quatro últimas semanas, que o modelo deve prever.

Esse formato simula o cenário operacional em que previsões são geradas para períodos futuros ainda não observados.

O modelo utilizado possui arquitetura enxuta, adequada para séries univariadas e para experimentos com janela curta. Ele é composto por:

- Uma camada *LSTM* com 50 neurônios e ativação *tanh*;
- Uma camada *Dense* (1 neurônio) para retornar a previsão.

A rede é treinada com:

- Otimizador *Adam* (learning rate = 0.01);
- Função de perda *MSE*.

Essa escolha busca equilibrar velocidade de treinamento e estabilidade da convergência.

Após o treinamento inicial com o conjunto de treino, foi aplicada a técnica de *rolling forecasting*, que simula o uso do modelo em produção. O procedimento funciona da seguinte forma:

1. Para cada ponto do conjunto de teste, o modelo recebe como entrada as últimas 3 semanas disponíveis;
2. O modelo produz uma previsão normalizada para a próxima semana;
3. A previsão é armazenada e, em seguida, o valor real observado naquela semana é adicionado ao histórico;
4. O modelo é reentrenado rapidamente com o histórico atualizado, utilizando poucas épocas (10), de forma a incorporar informações recentes.

Esse processo é repetido até que todas as semanas do período de teste tenham sido previstas. A abordagem *rolling forecasting* permite que o modelo esteja sempre adaptado às mudanças mais recentes da série, característica fundamental para ambientes hospitalares, onde o consumo pode variar rapidamente.

Como a série foi previamente normalizada para o intervalo [0, 1], as previsões geradas pela *LSTM* precisaram ser reconvertidas para a escala original utilizando o inverso do *Min-Max Scaler*. Isso permite que as previsões sejam comparadas diretamente com os valores reais da variável de interesse.

3. Resultados

3.1. Comparação dos modelos

Para a análise de desempenho dos modelos de previsão, foi escolhido o produto *Luva de Procedimento G*, responsável por aproximadamente 50% do custo total dos insumos de limpeza do hospital. Além disso, trata-se do item com maior consumo, o que torna a análise de sua previsão especialmente relevante para a gestão

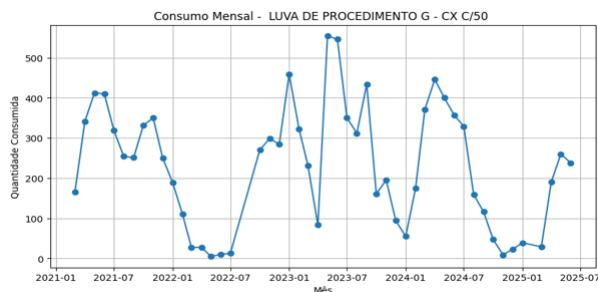


Figura 1: Consumo da Luva de Procedimento G ao longo do período analisado.

de estoques e planejamento de compras, conforme ilustrado na Figura 1.

Os modelos avaliados foram: *LSTM*, *ARIMA/SARIMA*, Média Móvel e *Rolling Forecasting* baseado em *LSTM*. Os resultados obtidos estão resumidos na Tabela 1.

Modelo	MSE	RMSE	MAE	R ² /MAPE
LSTM	19012.12	137.88	114.83	0.17
ARIMA/SARIMA	19012.12	41.97	-	250 (MAPE)
Média Móvel	14725.85	121.35	101.35	0.31
Rolling Forecasting (LSTM)	-	33.33	28.52	-

Tabela 1: Desempenho dos modelos de previsão para o produto *Luva de Procedimento G*.

As métricas apresentadas na Tabela 1 foram utilizadas para avaliar o desempenho preditivo dos modelos analisados. O *Mean Squared Error* (MSE) mede a média dos quadrados dos erros entre os valores observados e previstos, penalizando de forma mais severa erros de maior magnitude. O *Root Mean Squared Error* (RMSE) corresponde à raiz quadrada do MSE, expressando o erro médio na mesma unidade da variável original, o que facilita sua interpretação prática. O *Mean Absolute Error* (MAE) representa a média dos valores absolutos dos erros de previsão, sendo menos sensível a valores extremos quando comparado ao MSE e ao RMSE.

Além dessas métricas, o coeficiente de determinação (R^2) indica a proporção da variabilidade dos dados observados explicada pelo modelo, assumindo valores entre 0 e 1, sendo que valores mais próximos de 1 indicam melhor ajuste. Para o modelo ARIMA/SARIMA, optou-se pela utilização do *Mean Absolute Percentage Error* (MAPE), que expressa o erro médio em termos percentuais, permitindo avaliar a magnitude relativa dos erros de previsão em relação aos valores reais. Ressalta-se, entretanto, que os modelos foram treinados sob condições distintas, com diferentes configurações de parâmetros e estratégias de ajuste, visando extrair o melhor desempenho possível de cada abordagem, o

que deve ser considerado na interpretação comparativa dos resultados.

Os gráficos apresentados a seguir evidenciam visualmente a diferença de performance entre os modelos mencionados.

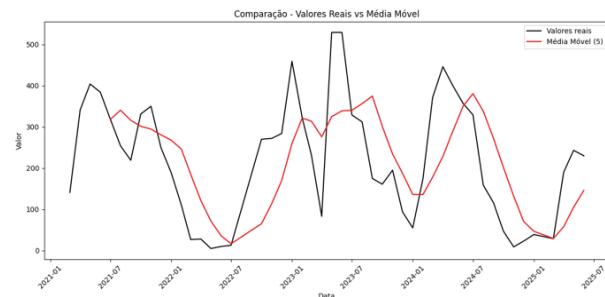


Figura 2: Performance da média móvel

A Figura 2 apresenta a comparação entre os valores reais de consumo e os valores estimados pelo modelo de Média Móvel com janela de cinco períodos. A série representada pela linha contínua corresponde aos valores observados, os quais apresentam elevada variabilidade ao longo do tempo, com oscilações abruptas e picos de consumo em determinados meses.



Figura 3: Performance dos modelos ARIMA e SARIMA

A Figura 3 apresenta a validação das previsões obtidas pelos modelos ARIMA e SARIMA, comparando os valores reais observados com as estimativas geradas para o período de teste. A série em cinza representa os dados utilizados no treinamento, enquanto a linha preta corresponde aos valores reais no intervalo de validação.

Observa-se que ambos os modelos produziram previsões praticamente constantes ao longo do horizonte analisado, conforme ilustrado pelas curvas referentes ao ARIMA e ao SARIMA. Esse comportamento indica que os modelos capturaram predominantemente o nível médio da série histórica, apresentando dificuldade em reproduzir a variabilidade e as oscilações observadas nos dados reais.

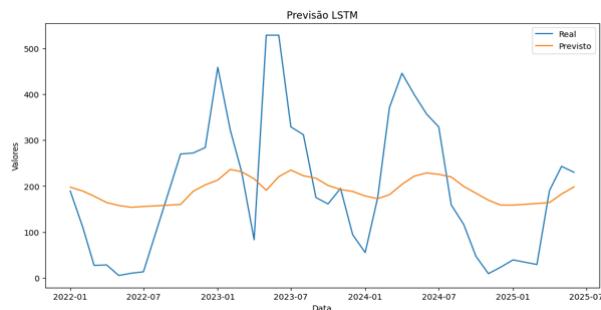


Figura 4: Performance da RNN LSTM

A Figura 4 apresenta a comparação entre os valores reais de consumo e os valores estimados pelo modelo de LSTM. A série representada pela linha contínua corresponde aos valores observados, os quais apresentam elevada variabilidade ao longo do tempo, com oscilações abruptas e picos de consumo em determinados meses. Pode ser notado que o modelo simples de LSTM captou pouco a variação dos dados reais.

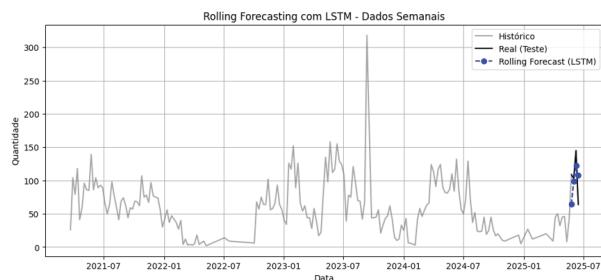


Figura 5: Performance do modelo rolling forecasting com LSTM

A Figura 5 apresenta a comparação entre os valores reais de consumo e os valores estimados pelo modelo de LSTM com aplicação do Rolling Forecasting. É possível notar que o modelo pode captar melhor a variação para o período a frente previsto. Fato este que já havia sido analisado na tabela 1

4. Conclusão

Entre todos os modelos avaliados, o LSTM com *Rolling Forecasting* apresentou o melhor desempenho, alcançando os menores valores de erro ($MAE = 28.52$ e $RMSE = 33.33$). Esse resultado demonstra a capacidade do modelo de se adaptar continuamente ao comportamento recente da série, reduzindo efeitos de *drift* e lidando melhor com a volatilidade observada no consumo do produto.

O modelo LSTM estático, por outro lado, apresentou ajuste limitado e erros significativamente mais altos, evidenciando dificuldade em capturar a dinâmica tem-

poral quando treinado apenas uma vez. Já os modelos ARIMA e SARIMA, mesmo após a seleção automática de parâmetros, mostraram desempenho insatisfatório, com MAPE elevado e incapacidade de acompanhar mudanças abruptas.

A Média Móvel serviu como *baseline* para o estudo, oferecendo uma referência simples de comparação, porém com precisão inferior às abordagens baseadas em redes neurais.

Assim, conclui-se que o uso de *Rolling Forecasting* com LSTM foi a estratégia mais adequada para a série temporal utilizada, apresentando melhor desempenho preditivo em comparação aos demais modelos avaliados. Ademais, vislumbra-se como possibilidade de trabalhos futuros a expansão do estudo por meio da aplicação do *Rolling Forecasting* em outros modelos de previsão, como ARIMA e SARIMA. Para trabalhos futuros, recomenda-se ainda aprofundar a comparação entre os métodos sob condições experimentais mais homogêneas, bem como avaliar o uso de abordagens híbridas e ajustes nas janelas temporais, de modo a ampliar a robustez e a confiabilidade das previsões.

Agradecimentos

Aos meus pais, cuja força, amor e presença silenciosa iluminaram cada passo desta jornada;

Aos meus amigos, pela companhia constante, pelos conversas que aliviaram os dias difíceis e pelas risadas que mantiveram o caminho leve;

A todos que contribuíram com dados, apoio e conhecimento para que este trabalho ganhasse forma;

A cada pessoa que um dia ousou seguir caminhos distintos, desafiando o óbvio e acreditando nas novas trajetórias.

A todos vocês, meu sincero agradecimento.

Referências

- [1] M. Martins, L. Monteiro, R. Silva e A. Costa. *Machine Learning with Environmental Predictors to Forecast Hospital Visits and Admissions: A Systematic Review*. Environmental Systems Research, Springer, (2025).
- [2] L. Tello, M. Tavakoli, A. Faghih-Roohi e S. Mansouri. *Care Process Optimization in a Cardiovascular Hospital: An Integration of Simulation–Optimization and Data Mining*. BMC Health Services Research, Springer, (2023).
- [3] A. Naser, M. Abou-Shouk, R. Elbanna e A. Mohamed. *Operational Flexibility Impact on Hospital Performance Through the Roles of Employee Engagement*

- and Management Capability.* Annals of Operations Research, Springer, (2023).
- [4] G. E. P. Box, G. M. Jenkins, G. C. Reinsel. *Time Series Analysis: Forecasting and Control.* 4. ed., Wiley, (2008).
 - [5] R. H. Shumway, D. S. Stoffer. *Time Series Analysis and Its Applications: With R Examples.* 4. ed., Springer, (2017).
 - [6] A. Inoue, L. Jin, B. Rossi. Rolling Window Selection for Out-of-Sample Forecasting with Time-Varying Parameters. *Journal of Econometrics*, v. 196, n. 1, pp. 55–67, (2017).
 - [7] K. Cho et al. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. *arXiv preprint arXiv:1406.1078*, (2014).
 - [8] S. Hochreiter, J. Schmidhuber. Long Short-Term Memory. *Neural Computation*, v. 9, n. 8, pp. 1735–1780, (1997).
 - [9] D. E. Rumelhart, G. E. Hinton, R. J. Williams. Learning representations by back-propagating errors. *Nature*, v. 323, pp. 533–536, (1986).
 - [10] A. Vaswani et al. Attention Is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*, (2017).