

## **1. Introdução**

A ArcelorMittal Brasil recebe semanalmente milhares de caminhões em suas plantas (mineração, fazendas, usinas, fábricas e distribuição). Ter uma previsão confiável de quantos veículos poderão chegar em um dado horizonte de tempo é fator crítico para viabilizar a eficiência da operação logística, permitindo alocar recursos de forma adequada (pessoas, máquinas e equipamentos), evitar custos não previstos (diárias, horas extras) e melhorar a satisfação do público envolvido, seja interno (equipes de portaria, de produção e funcionários em geral) ou externo (motoristas, fornecedores e comunidade local).

Para tratar esse problema foi desenvolvido um sistema de agendamento onde transportadoras agendam os veículos com base em uma data/hora prevista para chegada. A partir do agendamento, a Logística da planta industrial passa a ter visibilidade da quantidade de veículos prevista por perfil de operação (expedição ou recebimento de matéria-prima). Apesar das vantagens proporcionadas pelo sistema, a principal dificuldade é que a previsão da chegada dos veículos ocorre em um horizonte bastante curto (normalmente entre 1 a 3 dias).

O presente artigo é baseado em uma pesquisa interna com objetivo de validar a hipótese de que utilizar técnicas de previsão de séries temporais podem ampliar a visibilidade proporcionada pelo sistema, considerando horizontes mais amplos e menor erro de previsão.

Apesar das diferenças entre as abordagens utilizadas (detalhadas ao longo do trabalho), 2 dos 4 modelos avaliados proporcionaram erro de previsão 56% abaixo da referência utilizada e com horizonte de visibilidade superior.

Na seção 2 é realizada uma revisão da literatura sobre séries temporais e as técnicas escolhidas, na seção 3 é apresentada a metodologia de avaliação, seguida da apresentação dos dados e aplicação nos modelos e nas sessões 4 e 5 são apresentados os resultados e discussão.

## **2. Referencial teórico**

A previsão dos níveis de demanda é vital para a empresa como um todo, à medida que proporciona a entrada básica para o planejamento e controle de todas as áreas funcionais. Os níveis de demanda e os momentos em que ocorrem afetam fundamentalmente os índices de capacidade, as necessidades financeiras e a estrutura geral de qualquer negócio. A previsão logística abrange tanto a natureza espacial quanto a natureza temporal da demanda, a extensão

de sua variabilidade e seu grau de aleatoriedade. Muitos métodos de previsão de curto prazo trabalham com esse tipo de variação temporal, nas chamadas de séries temporais. Vários são os métodos de previsão disponíveis os quais são divididos em três categorias: qualitativos, de projeção histórica e causais (BALLOU, 2006).

Séries Temporais são observações feitas sobre um fenômeno ao longo do tempo. De acordo com Hyndman e Athanasopoulos (2018) os principais padrões de uma série temporal são: tendência, que caracteriza aumentos ou reduções a longo prazo nos dados; sazonalidade, que é quando a série é afetada por fatores sazonais como por exemplo valores mais altos no início do dia e mais baixos no fim do dia; e ciclos, que podem ser demonstrados como aumentos ou reduções sem período de tempo fixo.

Segundo Parmezan, Souza e Batista (2019) cada método possui vantagens e desvantagens, não havendo portanto, um único método capaz de resolver todo tipo de problema. Esse trabalho avalia 4 técnicas de previsão, com diferentes níveis de esforço e complexidade.

A média Móvel é baseada na ideia de que a demanda futura é similar a demanda recente observada. Com esse modelo nós simplesmente assumimos que a previsão é a média da demanda durante os últimos  $n$  períodos. Em contrapartida, se verificarmos a demanda em uma base mensal, podemos traduzir como "A previsão de demanda em junho será a média entre março, abril e maio" (VANDERPUT, 2018).

Já as técnicas de Suavização Exponencial decompõem séries temporais em componentes os quais seus valores são suavizados por pesos que caem ao longo do tempo. De acordo com Vanderput (2018), esses métodos frequentemente chamados de "Holt-Winters", em referência aos nomes dos pesquisadores que propuseram esses modelos. Uma forma inicial de previsão por suavização exponencial foi proposta por R.G. Brown em 1956 na qual as equações foram refinadas pelo trabalho de Holt (1957) e 03 anos mais tarde, os modelos de suavização exponencial foram novamente melhorados dessa vez por Winters (1960). Ambos nomes foram lembrados e atribuídos para todas diferentes técnicas que muitas das vezes são chamadas de "Holt-Winters".

Parmezan, Souza e Batista (2019) comentam que modelos auto-regressivos como o *ARIMA* (*Autoregressive Integrated Moving Average*) implicam na execução de 3 procedimentos estatísticos como entender como os valores atuais influenciam os próximos, garantir a estacionariedade da série e a utilização de média móvel para tratar fatores desconhecidos. A

técnica *SARIMA* (*Seasonal Autoregressive Integrated Moving Average*) atua da mesma forma que o *ARIMA* considerando parâmetros de sazonalidade.

O Aprendizado de Máquina (*Machine Learning*) é um termo que engloba diversas técnicas que proporcionam desenvolver um modelo preditivo. De acordo com Mitchell (1997) diz-se que um programa de computador aprende com a experiência *E*, em relação a uma classe de tarefas *T* e performance *P*, e se somente se sua performance nas tarefas *T* e performance *P*, melhora com a experiência *E*. Para Ahmed, Atiya, Gayar e El-Shishiny (2010) o aprendizado de máquina usa a teoria da estatística na construção de modelos matemáticos, porque a tarefa principal é inferência de uma amostra.

Uma vertente do Aprendizado de Máquina é o Aprendizado Profundo (*Deep Learning*) que utiliza algoritmos e arquiteturas de redes neurais para aprendizagem em dados não estruturados como imagens e texto. Redes Neurais Recorrentes (*Recurrent Neural Network – RNN*) é uma arquitetura de rede neural que permite trabalhar com dados temporalmente sequenciais. Em uma *RNN*, a rede com um input  $X_t$  e proporciona uma saída  $H_t$ . Essa característica encadeada permite com que uma informação seja passada de forma sucessiva.

Segundo Olah, (2015) Um dos problemas das *RNNs* é lidar com dependências de longo prazo, onde a rede pode deixar de propagar uma informação vários passos a frente. Para resolver esse problema Hochreiter e Schmidhuber (1997) propuseram a *LSTM* (*Long-Short Term Memory*), um tipo de *RNN* capaz de aprender dependências de longo prazo. Apesar de ser popular na literatura modelos *LSTM* são considerados complexos, de difícil treino e elevado custo computacional.

Graças ao aumento do poder computacional e disponibilidade de dados e ferramentas gratuitas, qualquer pessoa com habilidades de programação e disponibilidade para experimentação pode fazer diferença utilizando modelos de *machine learning* e proporcionar a qualquer negócio previsões melhores do que qualquer software de previsão de prateleira. Assim como os modelos estão se tornando melhores e mais rápidos, as ferramentas para utilizá-los estão se tornando mais amigáveis e fáceis de serem utilizadas (VANDERPUT, 2018).

Essas ferramentas são em sua maioria, estruturadas em bibliotecas computacionais, normalmente de código aberto, mantidas pela comunidade de software e empresas e que fornecem a implementação de diversas técnicas, algoritmos e funções comuns ao desenvolvimento de modelos de inteligência artificial.

Startups e empresas de tecnologia, principalmente as plataformas de motoristas de aplicativo como Uber, Lyft e Di-Di, tem se tornado referência no mercado e academia resolvendo problemas logísticos complexos com modelos estado da arte utilizando técnicas estatísticas e de *machine learning* em publicações como Laptev, Yosinski, Li e Smyl (2017) e Zhu, Laptev (2017).

Considerando as informações apresentadas, o presente trabalho propõe comparar as técnicas de previsão citadas com a previsibilidade disponível atualmente no sistema. Na seção seguinte serão apresentados os dados, critérios de avaliação e comparação e resultados de cada modelo em uma pequena amostra.

### 3. Metodologia e desenvolvimento

Na seção 3.1 é apresentada a estratégia geral de avaliação, na seção 3.2 é apresentado o conjunto de dados (*dataset*) e na seção 3.3 são apresentados os modelos testados e resultados iniciais.

#### 3.1 – Avaliação dos modelos

A previsibilidade proporcionada pelo sistema de agendamento difere tecnicamente dos métodos de previsão. Os tópicos seguintes abordam as diferenças gerais para alinhar entendimento e critérios.

##### 3.1.1 – Diferenças básicas entre os modelos

Com o sistema é possível saber a quantidade de veículos agendados para uma data e nesse trabalho serão considerados agendamentos realizados com 01 e 02 dias de antecedência. O agendamento de um veículo é considerado como o compromisso entre o fornecedor ou transportadora e uma planta industrial para que o motorista chegue até uma data/hora registrada no sistema, e esse fato não impede atrasos ou ausências.

No caso da Média Móvel a previsão pode ser calculada em praticamente todo o *dataset*, tendo início a partir do fim do período (por exemplo, uma média móvel de 3 períodos pode ser calculada a partir da 3ª observação). Para os modelos Suavização Exponencial (Holt-Winters), *SARIMA* e *LSTM* a previsão deve ser considerada a partir do *dataset* de teste, já que no *dataset* de treino o modelo deverá aprender o comportamento dos dados.

O *dataset* de treino é utilizado para ajustar o modelo com base no ajuste de parâmetros. Ele irá permitir avaliar a acurácia do modelo para dados “não vistos” e para isso, a amostra de teste é separada e somente executada depois da execução do modelo no *dataset* de treino (VANDERPUT, 2018). A Figura 2 exemplifica essa divisão para séries temporais.

### 3.1.2 – Métricas de erro

Medidas de erro de previsão são aplicadas para avaliar a qualidade dos métodos de previsão. A fórmula disposta na equação 1 considera que o erro ( $e_i$ ) é a diferença entre o valor medido no tempo ( $y_i$ ) pelo valor previsto ( $\hat{y}_i$ ) no tempo  $i$ .

$$e_i = y_i - \hat{y}_i \quad (\text{Eq. 1})$$

Os principais tipos de medidas de erro são os erros absolutos e relativos e cada métrica pode ter vantagens e desvantagens. A métrica escolhida para comparação dos métodos é a *Root Mean Squared Error (RMSE)* – Erro Médio Quadrático que é calculado pela equação 2.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i^2)} = \sqrt{\text{mean}(e_i^2)} \quad (\text{Eq. 2})$$

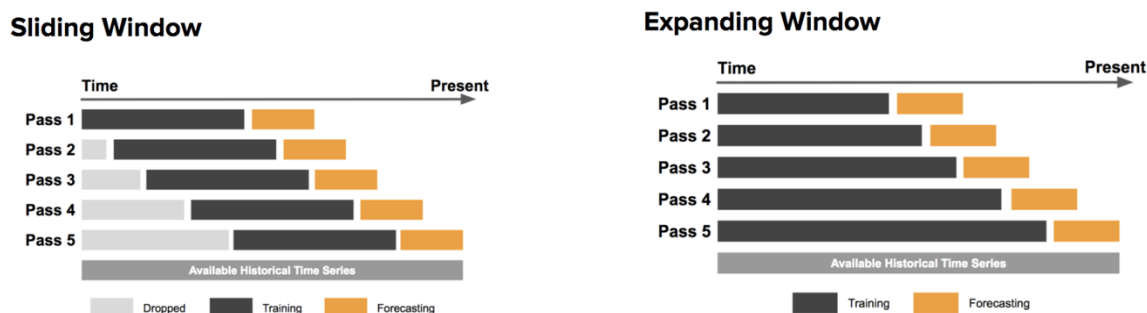
A principal vantagem do *RMSE* é ser sensível a grandes erros. Isso é importante pois grandes erros podem ser significativos no dimensionamento de recursos para recebimento dos caminhões.

### 3.1.3 – Amostras de teste

Para uma avaliação consistente de cada modelo é importante comparar os dados reais e de previsão em diferentes momentos da série. Isso torna o modelo mais preparado para prever novos dados e consequentemente se tornar mais apto para execução em ambiente de produção. O processo de *Walk Forward Validation* prevê que sejam definidas amostras de teste sequenciais ao longo da série que podem ser feitas pelo método *Sliding Window* no qual as amostras de treino e teste são congeladas e deslocadas a um passo definido de observações e a *Expanding Window (EW)* em que é congelado o início do *dataset* de treino e o fim aumenta

gradualmente a cada passo. A figura 1, exibe um comparativo entre as opções e nesse trabalho será utilizado o método *EW* devido a possibilidade de aumentar a amostra de treino.

Figura 1 – Comparação de amostras de treino e teste para os métodos *SW* e *EW*



Fonte: Bell e Smyl (2018)

### 3.1.4 – Critério final de avaliação

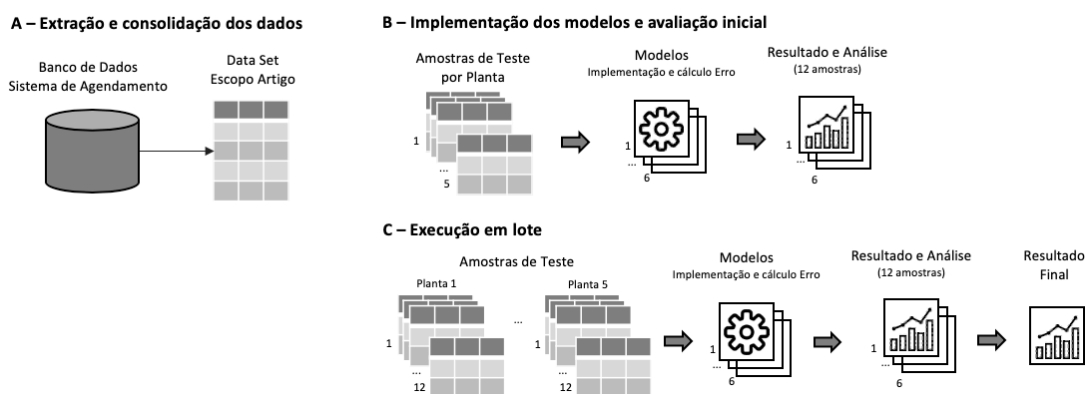
Os modelos serão comparados de 3 formas. A primeira na perspectiva do sistema de agendamento (comparando os diferentes horizontes de tempo disponíveis), a segunda comparando apenas as técnicas de previsão entre si, e a última comparando todos os modelos (independente das diferenças técnicas), já que cada um proporciona vantagens e desvantagens na solução do problema.

O *dataset* tem início em 01/01/2019 e fim em 31/12/2019. A amostra de treino tem início no período de 01/01/2019 a 01/09/2019 (8 meses) e a amostra de teste inicia no período de 02/09/2019 a 15/09/2019 (14 dias). A partir desse período é aplicada a técnica de *EW* em 12 passos sequenciais.

Será calculado o erro de previsão (*RMSE*) entre a quantidade de veículos recebidos e previstos para cada amostra de teste no período de 14 dias. Os modelos do Agendamento e Média Móvel não necessitam treinamento necessitando realizar apenas os testes.

A figura 2 resume o processo de desenvolvimento do modelo, onde os tópicos A e B são detalhados na seção 3.2 e 3.3 e os resultados obtidos no tópico C são apresentados na seção 4.

Figura 2 - Processo de desenvolvimento



Fonte: O presente trabalho

### 3.2 – O dataset

O *dataset* foi elaborado com os dados obtidos no próprio sistema de agendamento e foi modelado para apresentar os dados diários conforme apresentado na figura 3. Os dados referem-se aos veículos e plantas considerados no escopo desse trabalho e conta com as seguintes informações:

- Planta: Corresponde a uma das 5 plantas industriais onde os dados foram obtidos;
- Data: Data de referência dos registros (de 01/01/2019 a 31/12/2019);
- Recebidos: Quantidade de veículos recebidos no dia;
- Agendados (01 dia antes): Quantidade de veículos agendados para a data de referência com 01 dia de antecedência;
- Agendados (02 dias antes): Quantidade de veículos agendados para a data de referência com 02 dias de antecedência;

Figura 3 – Primeiras linhas do *dataset* (dados diários de 5 plantas no período de 01/01/2019 a 31/12/2019)

	Planta	Data	Recebidos	Agendados (01 dia antes)	Agendados (02 dias antes)
0	Planta_01	2019-01-01	0.0	0.0	0.0
1	Planta_01	2019-01-02	133.0	14.0	14.0
2	Planta_01	2019-01-03	97.0	61.0	4.0
3	Planta_01	2019-01-04	106.0	40.0	5.0
4	Planta_01	2019-01-05	38.0	30.0	4.0

No dia 05/01/2019, na planta 01 foram recebidos 38 veículos enquanto 30 foram agendados com um dia de antecedência e desse total, apenas 4 foram agendados com 2 dias de antecedência.

Fonte: Sistema de agendamento - ArcelorMittal Brasil.

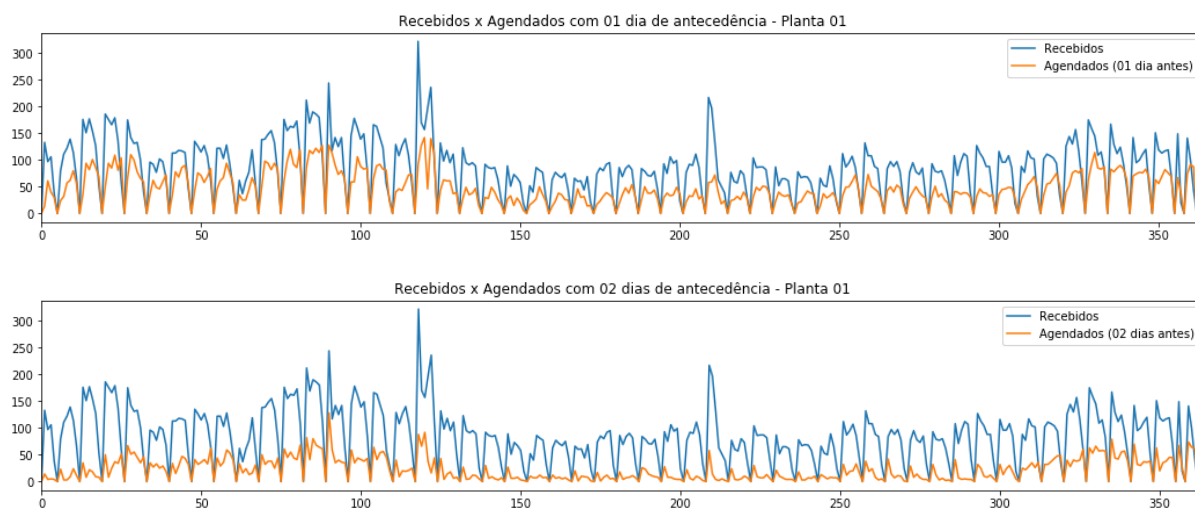
### 3.3 – Modelos de previsão

Os modelos foram desenvolvidos de acordo com a referência teórica de cada técnica nos quais os 3 primeiros foram calculados diretamente e os demais foram construídos utilizando bibliotecas específicas. As sessões 3.3.1 a 3.3.6 apresentam o resultado de cada técnica considerando a amostra inicial de avaliação (primeiro passo de teste - 02/09/2019 a 15/09/2019 para a Planta 01).

#### 3.3.1 – Previsibilidade do sistema de agendamento

A previsibilidade do sistema já está disponível no *dataset* e é apresentada na figura 4.

Figura 4 - Veículos recebidos e previsão sistema agendamento  
 para 01 (acima) e 02 dias de antecedência (abaixo) em 2019 para a planta 1



Fonte: Modelo desenvolvido pelo autor

Ao longo de 14 dias a previsibilidade obtida pelo sistema de agendamento apresenta um *RMSE* de 37,180 para uma previsão com 1 dia de antecedência e 57,233 para 02 dias de antecedência na Planta 1.

#### 3.3.2 – Média Móvel

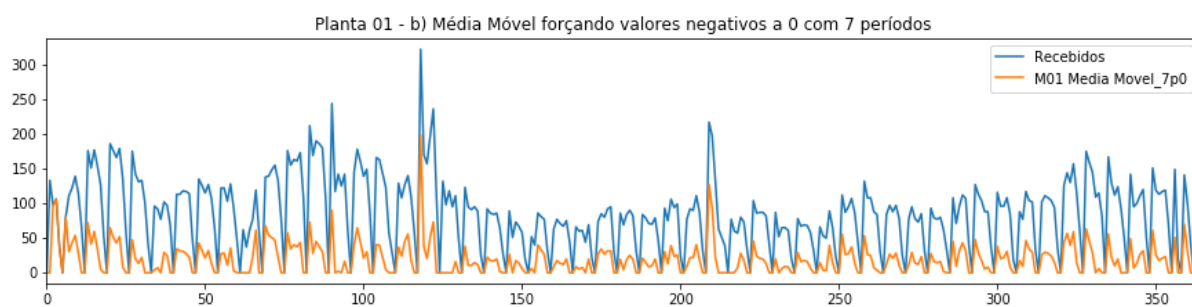
A primeira técnica de previsão utilizada é a Média Móvel. Por se tratar de um método bastante simples comparado aos demais, será considerada neste trabalho como uma abordagem *naive* (ingênua).



Um problema da média móvel aplicada aos dados utilizados é que pode ocorrer o cálculo de valores negativos, o que não faz sentido para o processo avaliado. Para contornar esse problema, o método será chamado de Média Móvel Ajustada (*naive*) no qual para cada valor de previsão negativo será considerado 0.

Foram simulados 04 cenários considerando períodos de 3 e 7 dias com e sem o ajuste. Desses o cenário com menor erro ( $RMSE = 50,883$ ) foi o com média móvel em 7 períodos (semana) ajustado. Essa técnica será considerada daqui pra frente como “Média Móvel Ajustada (*naive*)” e está demonstrada na figura 5.

Figura 5 – Veículos recebidos e previsão média móvel ajustada (*naive*) em 2019 para a planta 1

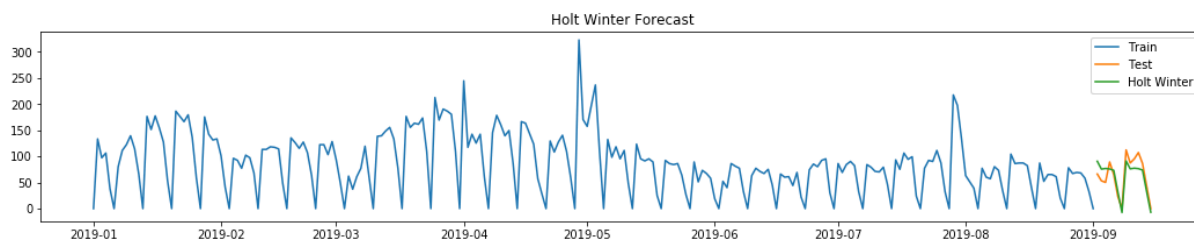


Fonte: Modelo desenvolvido pelo autor

### 3.3.3 – Suavização exponencial (Holt-Winters)

A implementação desse método foi realizada com a biblioteca *StatsModels*. Dentre os parâmetros necessários os principais são a definição de tendência e sazonalidade como “aditivos” já que esses valores não aumentam gradualmente ao longo das semanas e o período de sazonalidade indicado é de 7 dias (sazonalidade semanal). A figura 6 apresenta a série composta dos *datasets* de treino, teste e a previsão que para a mostra inicial apresentou um  $RMSE$  de 17,579.

Figura 6 – Veículos recebidos (01/01/2019 a 15/09/2019) e  
 previsão suavização exponencial (Holt-Winters) (02/09/2019 a 15/09/2019) para a planta 1

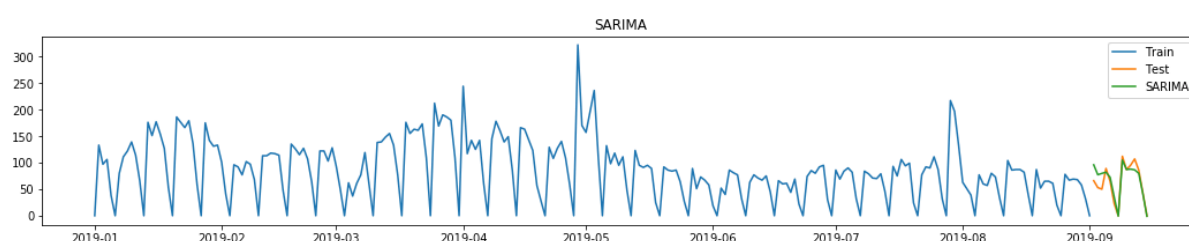


Fonte: Modelo desenvolvido pelo autor

### 3.3.4 – SARIMA

Essa técnica foi escolhida para avaliar se os modelos que consideram sazonalidade podem apresentar resultados muito diferentes entre si, principalmente nas 12 amostras de teste. Para a amostra inicial essa técnica apresentou *RMSE* de 15,114, sendo ligeiramente inferior a Suavização Exponencial. A alteração de parâmetros mais significativa foi considerar a média móvel e sazonalidade em 7 períodos ou dias e na figura 7 é possível observar que o modelo chega muito próximo ao *dataset* de teste.

Figura 7 – Veículos recebidos (01/01/2019 a 15/09/2019) e previsão *SARIMA* (02/09/2019 a 15/09/2019) para a planta 1

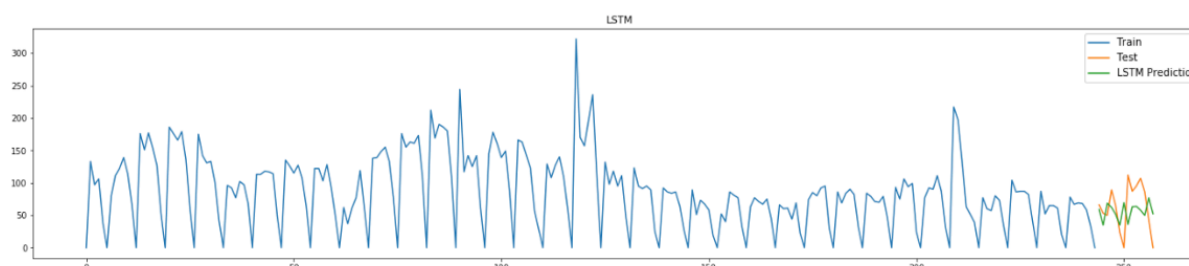


Fonte: Modelo desenvolvido pelo autor

### 3.3.5 – LSTM

Esse foi o modelo mais complexo em todo o trabalho. Os principais parâmetros considerados e que viabilizaram menor *RMSE* foi a definição de 100 camadas ocultas e 250 épocas de treinamento. Quanto maiores esses parâmetros, maior tende a ser o treinamento que para a amostra utilizada durou cerca de 5 horas que proporcionou a previsão apresentada na figura 8. Como a rede neural inicia sua execução com pesos aleatórios, os resultados podem variar a cada execução e o *RMSE* obtido para a amostra foi de 23,562.

Figura 7 – Veículos recebidos (01/01/2019 a 15/09/2019) e previsão *LSTM* (02/09/2019 a 15/09/2019) para a planta 1



Fonte: Modelo desenvolvido pelo autor

## 4. Resultados

A Tabela 1 apresenta um resumo dos resultados obtidos do *RMSE* após execução dos 6 modelos para as 12 amostras de teste em cada uma das 5 plantas.

Tabela 1 – RMSE médio das 12 amostras de teste por planta e o resultado final considerando a média do RMSE de cada planta ponderado pelos veículos recebidos

Planta	Qtd. Veículos	Abordagem					
		Agendamento (horizonte 1 e 2 dias)		Previsão Séries Temporais (horizonte 14 dias)			
		1 - Agendamento 01 dia	2 - Agendamento 02 dias	3 - Média Móvel Ajustada ( <i>Naive</i> )	4 - Suavização Exponencial (Holt-Winters)	5 - SARIMA	6 - LSTM
Planta 01	30.635	46,1	66,6	64,5	18,0	17,4	39,9
Planta 02	38.852	74,6	109,7	111,2	28,6	29,7	60,3
Planta 03	13.532	14,1	28,3	23,4	9,9	9,4	28,0
Planta 04	31.942	36,3	75,2	75,0	21,8	21,2	61,8
Planta 05	5.910	7,4	12,4	12,2	6,2	6,2	9,9
Resultado Final	-	47,2	75,8	75,2	20,9	20,9	49,4

Fonte: Resultado final do modelo desenvolvido pelo autor

As plantas com maior *RMSE* são também as que receberam o maior número de caminhões no período avaliado.

Comparando as abordagens do agendamento, a previsibilidade piora 61% entre o 1º e 2º dia. Dentre as 4 técnicas de previsão, a Suavização Exponencial (Holt-Winters) e SARIMA ficaram tecnicamente empatados, seguidos da LSTM e a Média Móvel Ajustada (*naive*).

Avaliando todas as abordagens (desconsiderando as diferenças técnicas entre si) é possível concluir que as técnicas de Suavização Exponencial (Holt-Winters) e SARIMA proporcionam uma previsibilidade com RMSE 56% abaixo do baseline (agendamento com 1 dia de antecedência) em um horizonte de previsão de 14.

## 5. Discussão

Foi mensurada a previsibilidade proporcionada pelo sistema de agendamento considerando em 2 horizontes de tempo (1 e 2 dias) e 4 técnicas de previsão de séries temporais no horizonte de 14 dias.

Os Modelos de Suavização Exponencial (Holt-Winters) e *SARIMA* apresentaram os melhores resultados, quando comparados com as demais técnicas de previsão e com o sistema.

A comparação das técnicas por amostragens diferentes se provou necessária, já que alguns modelos apresentaram relativa variabilidade ao longo das amostras de teste.

Conforme citado no próprio título desse trabalho o objetivo das descobertas é fazer com que as técnicas possam ampliar a previsão disponível no sistema, já que cada abordagem possui suas vantagens e desvantagens e uma complementa a outra.

Para trabalhos futuros são consideradas as seguintes estratégias:

- Utilizar técnicas avançadas para parametrização dos modelos como testes estatísticos e otimização;
- Incluir variáveis externas em modelos como SARIMAX e LSTM;
- Avaliar previsão por material e faixas de horário ao longo dos dias.

## 6. Referências

AHMED, Nesreen; ATIYA, Amir; GAYAR, Neamat; EL-SHISHINY, Hisham. **An Empirical Comparison of Machine Learning Models for Time Series Forecasting**. 2010. *Econometric Reviews*. 29. 594-621. 10.1080/07474938.2010.481556.

BALLOU, Ronald H. **Gerenciamento da cadeia de suprimentos/logística empresarial** / Ronald H. Ballou ; tradução Raul Rubenich. – 5. Ed. – Porto Alegre : Bookman, 2006

BELL, Franziska; SMYL, Slawek, **Forecasting at Uber: An Introduction**. 2018. Disponível em <<https://eng.uber.com/forecasting-introduction/>> Acesso em 15/05/2020.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. 1997. **Long Short-term Memory**. *Neural computation*. 9.1735-80. 10.1162/neco.1997.9.8.1735.

HOLT, Charles C.. **Forecasting seasonals and trends by exponentially weighted moving averages**. *International Journal of Forecasting* 20.1 (2004), pp. 5-10 (reprint original 1957)

HYNDMAN, Rob J.; ATHANASOPOULOS, George. **Forecasting: principles and practice**. 2018. 2nd edition, OTexts: Melbourne, Australia. Disponível em: <<https://otexts.com/fpp2/>> Acesso em 18/05/2020.

LAPTEV, Nikolay; YOSINSKI, Jason; LI Li Erran Li; SMYL, Slawek. **Time-series extreme event forecasting with neural networks at Uber**. 2017. In Int. Conf. on Machine Learning Time Series Workshop, 2017.

MITCHELL, Tom. **Machine Learning**. 1997. McGraw Hill, (Março, 1997).

OLAH, Christopher, **Understanding LSTM Networks**. 2015. Disponível em:  
<<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>> Acesso em 14/05/2020.

PARMEZAN, Antônio Rafael Sabino; SOUZA, Vinícius Alves de; BATISTA, Gustavo. **Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model**. 2019. Information Sciences. 10.1016/j.ins.2019.01.076.

VANDERPUT, Nicolas. **Data Science for Supply Chain Forecast**. 2018. Independently published (November 12, 2018)

WINTERS, Peter R.. **Forecasting Sales by Exponentially Weighted Moving Averages**. Management Science 6.3 (1960). Pp. 324-342. URL Disponível em: <<https://doi.org/10.1287/mnsc.6.3.324>>

ZHU, Lingxue; LAPTEV, Nikolay. **Deep and confident prediction for time series at Uber**. 2017. In Data Mining Workshops (ICDMW), IEEE International Conference on, pages 103–110. IEEE, 2017.

## 7. Anexos

Código fonte do Modelo disponível no GitHub em:

<https://github.com/aburachid/Time-Series-ENEGEP-Paper>