



UNIVERSIDADE FEDERAL DO PARANÁ

VINICIUS STRUGATA AMBROSIO

KOJI SASAKURA

PROPOSTA DE UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA PARA PREDIZER A
DEMANDA FUTURA DE PEÇAS DE REPOSIÇÃO EM CONCESSIONÁRIAS DE
VEÍCULOS

CURITIBA

2019

VINICIUS STRUGATA AMBROSIO
KOJI SASAKURA

PROPOSTA DE UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA PARA PREDIZER A
DEMANDA FUTURA DE PEÇAS DE REPOSIÇÃO EM CONCESSIONÁRIAS DE
VEÍCULOS

Trabalho de Conclusão de Curso apresentado ao curso de Pós-Graduação em Engenharia Industrial 4.0, Setor de Tecnologia, Universidade Federal do Paraná, como requisito à obtenção do título de Especialista em Engenharia Industrial 4.0.

Orientador: Prof. Dr. Pablo Deivid Valle

CURITIBA

2019

TERMO DE APROVAÇÃO

VINICIUS STRUGATA AMBROSIO
KOJI SASAKURA

PROPOSTA DE UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA PARA PREDIZER A DEMANDA FUTURA DE PEÇAS DE REPOSIÇÃO EM CONCESSIONÁRIAS DE VEÍCULOS

Trabalho de Conclusão de Curso apresentado ao curso de Pós-Graduação em Engenharia Industrial 4.0, Setor de Tecnologia, Universidade Federal do Paraná, como requisito à obtenção do título de Especialista em Engenharia Industrial 4.0.

Prof. Dr. Pablo Deivid Valle

Orientador(a) – Departamento de Engenharia Mecânica, UFPR

Curitiba, 19 de outubro de 2019.

RESUMO

Este trabalho tem como tema principal apresentar uma proposta de utilização de um algoritmo de aprendizado de máquina na gestão enxuta do estoque de itens de reposição de uma concessionária de veículos. Este algoritmo utiliza a coleta e análise de dados para aprimorar, de modo contínuo e automatizado, a previsão da compra dos itens de estoque, otimizando assim a gestão de compras, mas sem prejudicar o compromisso com o nível de serviço ao cliente. Deste modo a redução dos custos de estoque ocorrerá através de uma compra inteligente, na medida certa das necessidades do mercado, porém sem a necessidade de análises manuais e decisões feitas com base na intuição do gestor de estoque.

Palavras chave: Aprendizado de Máquina. Custos de estoque. Gestão Enxuta. Algoritmo. Coleta de dados. Demanda Futura.

ABSTRACT

This main goal of this paper is to present a proposition for the usage of a machine learning algorithm in the lean management of the stock of replacement items from a vehicle dealership. This algorithm uses data collection and analysis to improve, in a continuous and automated way, the inventory items purchase prediction, thus optimizing the management of purchases, but without jeopardizing the commitment to the level of customer service. In this way the reduction of inventory costs will occur through an intelligent purchase, in the right measure of the market needs, but without the need for manual analyzes and decisions made based on the intuition of the inventory manager.

Keywords: Machine Learning. Inventory costs. Lean Management. Algorithm. Data collect. Future Demand.

LISTA DE ILUSTRAÇÕES

FIGURA 1 - EXEMPLO DE UM NEURÔNIO.....	14
FIGURA 2 - EXEMPLO DE UMA REDE NEURAL PROFUNDA.....	15
FIGURA 3- EXEMPLO DE COMO O MÉTODO DO GRADIENTE FUNCIONA. O "CAMINHO" SEGUIDO PELO MÉTODO ENCONTROU UM PONDO DE MINIMO LOCAL, EMBORA EXISTA OUTRO MINIMO GLOBAL (OBSERVAR AS INDICAÇÕES).....	16
FIGURA 4 - EXEMPLO DE UMA ARVORE DE DECISÃO	18
FIGURA 5 - EXEMPLO DE COMO A MÁQUINA DE VETORES DE SUPORTE FUNCIONA.....	19
FIGURA 6 - GRÁFICO DE SETORES MOSTRANDO O PERFIL DO ESTOQUE NA CONCESSIONARIA AVALIADA NESTE ESTUDO. NOTA-SE QUE 26% DOS ITENS ESTÃO ACIMA DE 12 MESES DO PERÍODO DE COBERTURA.....	21
FIGURA 7 - UM CARRO PODE SER COMPOSTO DE MILHARES DE PEÇAS	23
FIGURA 8 - ESQUEMA MOSTRANDO COMO FUNCIONARIA A PROPOSTA DE SOLUÇÃO PARA O PROBLEMA IDENTIFICADO NA CONCESSIONARIA AVALIADA	26
FIGURA 9- TABELA COM UM EXEMPLO DE COMO SERIAM OS DADOS DO AMBIENTE NO QUAL O VEÍCULO PODE ESTAR CIRCULANDO	27
FIGURA 10 - TABELA COM UM EXEMPLO DE COMO SERIAM OS DADOS COLETADOS ATRAVÉS DOS SENSORES DOS VEÍCULOS.....	27
FIGURA 11 - EXEMPLO ILUSTRANDO A INFLUÊNCIA DA REGIÃO DE CIRCULAÇÃO NAS PEÇAS DE UM VEÍCULO.....	28
FIGURA 12- POTENCIAIS CLIENTES DA SOLUÇÃO A SER DESENVOLVIDA	30
FIGURA 13 - FASES DE IMPLANTAÇÃO DA SOLUÇÃO	33
FIGURA 14 - INVESTIMENTO NECESSÁRIO PARA IMPLANTAR A SOLUÇÃO, CONSIDERANDO UM PERÍODO DE 24 MESES	34
FIGURA 15 - PLACA GRÁFICA GPU USADA EM ALGORITMOS DE APRENDIZADO DE MÁQUINA.....	36

SUMÁRIO

1	INTRODUÇÃO	8
1.1	CONTEXTO DO PROBLEMA	8
1.2	HIPÓTESE DE SOLUÇÃO	8
1.3	OBJETIVOS E RESULTADOS ESPERADOS	9
2	REVISÃO BIBLIOGRÁFICA	10
2.1	BIG DATA.....	10
2.2	ESTRATÉGIAS	11
2.3	APRENDIZADO DE MÁQUINA.....	12
2.4	MODELOS	13
2.4.1	REDES NEURAIS ARTIFICIAIS	13
2.4.2	ÁRVORES DE DECISÃO	17
2.4.3	MÁQUINAS DE VETORES DE SUPORTE	18
3	CENÁRIO ATUAL	20
4	PROPOSTA DE SOLUÇÃO	23
4.1	IMPLEMENTAÇÃO	26
4.2	PÚBLICO ALVO E FORMA DE ENTREGA DO PRODUTO.....	29
4.3	INVESTIMENTO NECESSÁRIO E ESTRATÉGIA DE COMERCIALIZAÇÃO	31
5	CONCLUSÃO	36

1 INTRODUÇÃO

O estudo do aprendizado de máquina, ou Machine Learning na terminologia em inglês, iniciou-se com o pioneiro da inteligência artificial, Arthur Samuel (1959), engenheiro do MIT. Foi ele quem criou o termo aprendizado de máquina, descrevendo o conceito como um campo de estudo que dá aos computadores a habilidade de aprender sem terem sido programados para tal. Naquela época, Samuel trabalhava em um projeto para criar uma máquina autônoma com estas características.

1.1 CONTEXTO DO PROBLEMA

O presente trabalho será limitado ao estudo de utilização de algoritmo de aprendizado de máquina aplicado para otimização do custo de estoque de uma concessionária de veículos. Nesse contexto, irá ser tratada a dificuldade de manter os custos de estoque sob controle, de preferência o mais enxuto possível, porém sem comprometer os níveis de atendimento do cliente, pois a concessionária tem metas de desempenho de nível de serviço ao cliente impostas pela montadora a qual está vinculada.

1.2 HIPÓTESE DE SOLUÇÃO

Uma solução óbvia para este problema seria a redução dos custos dos itens mais relevantes na listagem dos itens de estoque com maiores valores. Porém essa ação não necessariamente é efetiva, pois não considera o comportamento de diversos fatores tais como *lead time* de entrega dos materiais, preço de compra, demanda sazonal, meta de compras, dentre outros. Além disso, ela não é determinística, pois desconsidera a característica dinâmica do mercado com o passar do tempo:

- O nível de redução a ser atingido sem comprometer o nível de atendimento dificilmente será o mesmo durante o ano
- O conjunto de peças mais relevantes também não será o mesmo durante o ano

Desta forma, o gestor de estoque usará os resultados do algoritmo de aprendizado de máquina para obter uma maior assertividade na demanda futura de peças de reposição, visando reduzir os efeitos deste comportamento dinâmico e muitas vezes aleatório do mercado e assim mesmo conseguir manter os níveis de estoques enxutos.

Este algoritmo irá utilizar de milhares de dados coletados, obtidos através dos sistemas de gestão da empresa (CRM, ERP), de dados históricos do mercado de comércio de veículos e de autopeças, de dados climáticos e por fim através de dados coletados por sensores localizados nos veículos. Será através da ingestão desta imensa quantidade de dados que o algoritmo terá a capacidade de identificar padrões e estabelecer correlações que serão apresentadas como recomendações de ações que permitirão a gestão enxuta do estoque de peças (compra de itens, promover ações incentivo a venda de itens etc.).

1.3 OBJETIVOS E RESULTADOS ESPERADOS

O objetivo geral para a solução deste problema será a redução dos custos de estoque utilizando um conjunto de regras determinadas pela concessionária no qual irão definir parâmetros ótimos para a compra/venda de itens de estoque. O algoritmo de aprendizado de máquina irá implementar modelos matemáticos que analisarão dados amplos e complexos de forma rápida e automática para entregar resultados precisos em grande escala.

Esses algoritmos usam análises estatísticas, identificam padrões e correlações para prever as respostas mais precisamente, entregando o melhor resultado com menor chance de erro. O resultado desta solução irá possibilitar custos reduzidos de estoque de acordo com as previsões de mercado e com a vantagem de que a máquina estará sempre se aprimorando, uma vez que a coleta e atualização dos dados ocorre de forma contínua e online. Ou seja, quanto maior o tempo em que o algoritmo de aprendizado de máquina estiver funcionando e coletando dados, mais assertivas serão suas respostas.

2 REVISÃO BIBLIOGRÁFICA

Dados estão em toda parte em uma organização, seja ela comercial, prestadora de serviços ou industrial. Taxa de serviços realizados por dia, quantidade de peças em estoque, esforço em homem-hora para esta produção e custos associados à produção. São apenas alguns exemplos de dados que um gestor poderia estar interessado.

A aquisição e interpretação dos dados desempenham um importante papel nas atividades de gestão em controle de produção ou de um estoque de peças. Uma interpretação inteligente destes permite que erros não sejam repetidos e fornece pistas para possíveis aperfeiçoamentos.

Com o aumento da digitalização dos processos industriais a partir da década de 60, tornou-se um desafio manter todos os dados de interesse sob controle. Não apenas isso, criou-se dificuldades para filtrar e validar tais dados de forma a não tirar conclusões a partir de dados corromptos ou não relevantes. Mais recentemente, tecnologias como *IoT – Internet of Things* (Internet das Coisas), com diversos sensores instalados em diferentes componentes ou locais, potencializaram a geração de dados, tornando assim este problema ainda mais proeminente.

Em informática, a área de estudos que trata dos desafios de se processar grandes quantidades de dados chama-se *Big Data*.

2.1 BIG DATA

“Big Data representa ativos de informação caracterizados por volume, velocidade e variedade tão altos que requerem tecnologia específica e métodos analíticos para sua transformação em valor” (De Mauro, *et al* (2016)). Esta definição foi aprimorada posteriormente para acrescentar mais uma característica relevantes: veracidade. Em resumo, tem-se então como conceitos – os cinco V’s – fundamentais do Big Data:

- Volume: dados são gerados em grande quantidade, requerem infraestrutura para armazenamento e processamento;
- Variedade: dados são de diferentes naturezas, estruturados ou não, aumentando em muito a complexidade de análise;

- Velocidade: dados são gerados a uma frequência muito elevada, e exigem grande velocidade de processamento, visto que as análises normalmente devem ser entregues em tempo real;
- Veracidade: dados podem estar incompletos, imprecisos ou não confiáveis. Deve-se usar de estratégias para garantir que as informações são verdadeiras e para o tratamento de dados incompletos;
- Valor: refere-se à utilidade e importância dos resultados obtidos através do processamento dos dados;

Percebe-se que cada uma destas características representa um grande desafio para o cientista de dados, nome dado a profissão responsável pela implantação dos conceitos do Big Data em uma organização. Tal profissional deve-se valer ainda de conhecimentos em estatística, banco de dados, programação de computadores, computação em nuvem, arquitetura de hardware de servidores, business intelligence e principalmente de comunicação e apresentação dos resultados das análises.

Questões de infraestrutura são extremamente relevantes e desafiadores quando o assunto é *Big Data*. De acordo com previsões do IDC, o volume de dados gerados globalmente irá chegar a 44 zettabytes em 2020, com previsão de 163 zettabytes em 2025. Igualmente com números astronômicos, consultoria Gartner aponta que em 2020 haverá aproximadamente 26 bilhões de dispositivos *IoT* conectados à internet e gerando dados. Tal quantidade de dispositivos conectados aumenta em muito preocupações com segurança da informação, principalmente em áreas como Energia, Saúde, Bancos/Finanças, Governo e Forças Armadas. Estimativas apontam uma projeção de cerca de 6 trilhões de dólares anuais em custos associados a crimes cibernéticos em 2021.

2.2 ESTRATÉGIAS

No tocante às estratégias para análise e tratamento dos dados, diversas técnicas vêm sendo desenvolvidas no decorrer das últimas décadas. Tradicionalmente, uma simples análise estatística descritiva básica (média, desvio padrão etc.) fornece informações valiosas sobre os dados. Tal análise é particularmente útil quando comparamos resultados em diferentes datas, gerando assim uma série temporal. Afinal, pode-se admitir que, se os fatores que produziram determinado resultado no passado se repetirem no futuro,

resultados semelhantes seriam obtidos. Técnicas probabilísticas, regressão linear e o uso de gráficos de controle poderiam ser adicionadas e assim aprimorar as análises.

Tal técnica funciona satisfatoriamente na maioria das situações. Entretanto, quando a análise necessita ser realizada para grandes volumes de dados de diferentes naturezas (estruturados, não estruturados), torna-se difícil conseguir resultados com a confiabilidade, precisão e agilidades requeridas. A estratégia ideal seria aquela capaz de ser genérica e rápida o suficiente para poder extrair resultados com as características desejadas anteriormente mencionadas, além de ser capaz de adaptar-se às variações dinâmicas dos dados mais recentes. Tal estratégia existe e chama-se Aprendizado de Máquina (Machine Learning).

2.3 APRENDIZADO DE MÁQUINA

No Aprendizado de Máquina, usa-se algoritmos e técnicas estatísticas para construir um modelo matemático capaz de identificar padrões presentes nos dados. Tais padrões são “ensinados” usando-se uma fração destes dados cuja resposta é conhecida, consistindo no que é chamado Dados de Treinamento. Por exemplo, um modelo pode ser treinado para identificar tumores de câncer de pele ao analisar a base de dados de imagens (e seus respectivos diagnósticos – os dados de treinamento) de um hospital. O final do treinamento ocorre quando a precisão desejada da identificação correta dos padrões é atingida.

Em seguida, este mesmo modelo matemático treinado é aplicado em outro conjunto de dados: os Dados de Teste. Nos dados de teste as respostas também são conhecidas, e o algoritmo apenas infere os resultados com base nos dados treinados. Novamente é avaliada a precisão da identificação correta das respostas. Caso a precisão das respostas for semelhante à dos dados de treinamento, significa que o modelo matemático está genérico o suficiente para ser utilizado em novos dados, desta vez com resultados desconhecidos. Ou seja, os dados de teste servem apenas para avaliar o quão preciso e genérico é o modelo em questão. Caso contrário, será necessário modificá-lo, ou aumentar a quantidade de dados de treinamento, até que a precisão das respostas seja atingida.

O que foi descrito anteriormente consiste no que se chama aprendizado supervisionado. Existem ainda diferentes abordagens:

- Aprendizado não supervisionado: o modelo matemático encontra por si só similaridades entre respostas;
- Aprendizado por reforço: consiste em uma técnica iterativa, com base no conceito de recompensas/penalizações, na qual o modelo matemático aprende com seus próprios erros e acertos. Assim, em uma próxima iteração, usa uma heurística para tentar conseguir uma recompensa maior ou uma penalidade menor. Exemplo de aplicação em jogos eletrônicos: encontrar a saída de um labirinto no menor tempo possível.

Quando o resultado desejado do modelo matemático é um número, diz-se que se trata de um problema de Regressão. Por exemplo, a previsão do preço de uma residência, a partir de dados históricos de uma imobiliária, usando sua localização, número de quartos e metragem, pode ser categorizada como um problema de regressão. Outra categorização bastante comum chama-se Classificação. A determinação da presença ou não de tumor de câncer de pele, citada anteriormente, é um problema de classificação.

2.4 MODELOS

Diferentes modelos matemáticos vêm sendo pesquisados para implementar o Aprendizado de Máquina. Nas seções a seguir, serão apresentados três dos modelos mais comumente usados:

2.4.1 REDES NEURAIIS ARTIFICIAIS

A proposta das RNAs (Redes Neurais Artificiais) consiste em conceber um algoritmo computacional usando técnicas matemáticas para estabelecer um paralelo ao funcionamento de um neurônio e suas iterações dentro do cérebro. A RNA recebe como entrada um conjunto de dados e devolve como resultado um número racional ou um conjunto de probabilidades, dependendo se a análise em questão for uma Regressão ou uma Classificação.

Como o próprio nome sugere, as RNAs são compostas por neurônios artificiais. Tais neurônios recebem um ou mais valores x_n de entrada, normalmente ponderados por um fator w_n em seus “dendritos”. Esses valores são somados antes de serem submetidos à uma função de ativação ou função de transferência φ . O resultado \hat{y} é propagado para o

próximo neurônio da rede, através de uma “sinapse”. O conjunto das iterações entre diferentes neurônios constitui a estrutura da RNA.

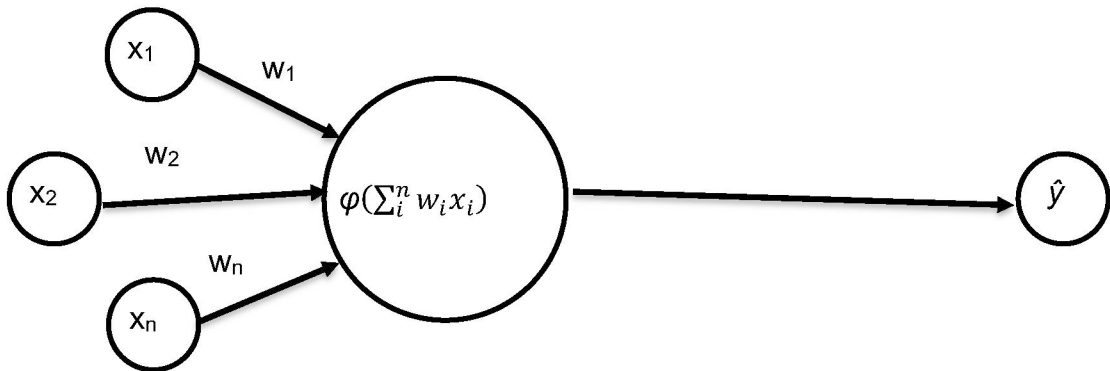


Figura 1 - Exemplo de um neurônio

Esta estrutura pode ser organizada em camadas, formando uma rede neural mais complexa, as chamadas Redes Neurais Profundas (RNP). Na figura abaixo, os neurônios designados com o prefixo h fazem parte das camadas ocultas. Aqueles com o prefixo i são os dados de entrada e os com o prefixo o são os resultados.

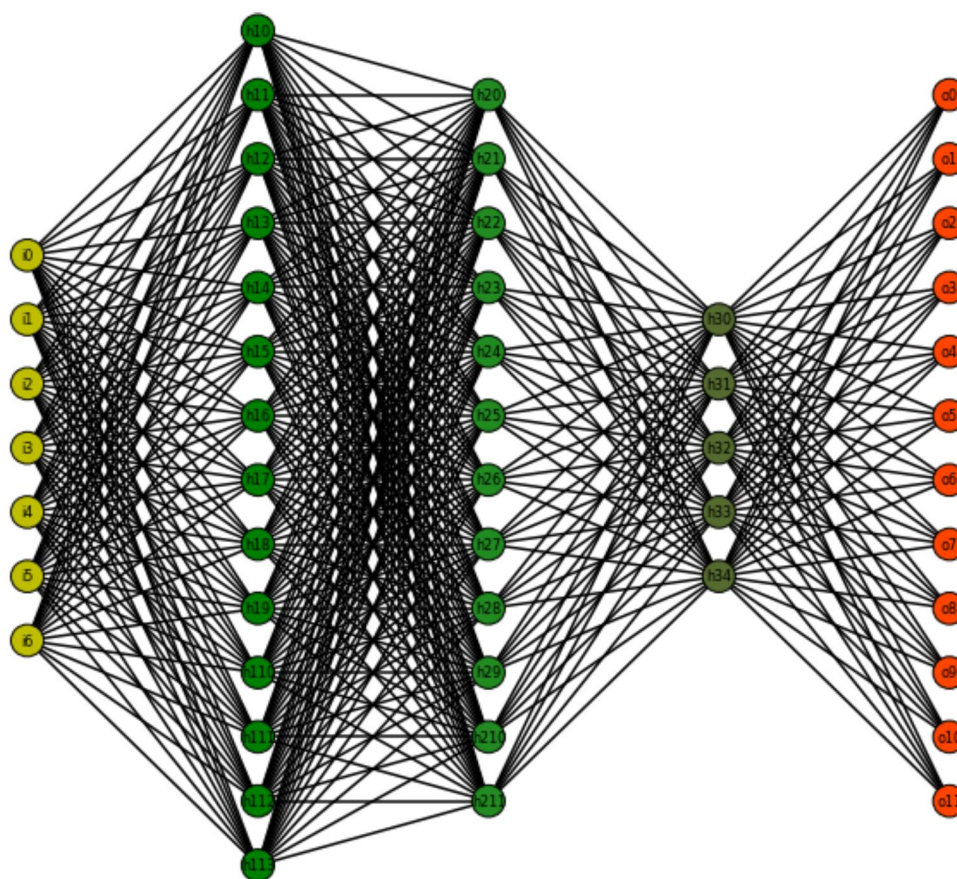


Figura 2 - Exemplo de uma Rede Neural Profunda

Cada camada oculta é responsável por aprender determinada característica ou padrão dos dados advindos da camada anterior. Por exemplo, em uma RNP para identificar objetos em imagens, pode-se ter camadas responsáveis por identificar bordas, outra responsável por identificar formas circulares etc. A quantidade e o tamanho das camadas estão relacionados com a capacidade de generalização da RNP. Ou seja, quanto mais camadas e quanto maior o número de neurônios, maior será a precisão na identificação de objetos.

Ainda na figura acima, cada neurônio é conectado com todos os neurônios da camada seguinte. Diz-se, neste caso, que se trata de uma estrutura RNP totalmente conectada. Neste tipo de estrutura a informação vinda dos dados de entrada acaba por “passar” por todos os neurônios da rede. Embora isso possa ser, a princípio desejável, visto que cada neurônio pode ter influência no resultado, na verdade pode propiciar que a RNP se torne sobre ajustado (*overfitted*). Quando isso acontece, a RNP tem um excelente desempenho nos dados de teste, mas é incapaz de manter performance similar em dados

inéditos. Idealmente, a RNP deveria possuir a mesma performance em dados inéditos daquela experimentada nos dados de teste e de validação.

Algumas técnicas ajudam a minimizar este inconveniente, tais como o desligamento aleatório de alguns neurônios das camadas ocultas durante o treinamento, regularização (penaliza a importância de determinados parâmetros de entrada que poderiam ser considerados ruídos) e *data-augmentation* (cria variações dos dados, a partir dos originais, aumentando assim o conjunto de dados processados nos casos em que estes são escassos).

Na RNP, inicialmente, os valores dos pesos w_n são atribuídos aleatoriamente. Durante a fase de treinamento, os valores dos pesos são otimizados com o objetivo de reduzir o erro dos resultados. Define-se como Época uma passagem completa dos dados de entrada na RNP. Ao fim de uma época, o erro é avaliado por intermédio de uma Função Erro e usa-se um Otimizador para recalculer os pesos visando reduzir o erro para a época seguinte. A função erro pode ser implementada usando-se a definição da fórmula do Erro Quadrático Médio. Com relação ao otimizador, as técnicas mais utilizadas envolvem a implementação de uma variação do Método do Gradiente (ou Método do Máximo Declive) para ajustar/recalculer os pesos em cada época. O processo finaliza quando o erro desejado é encontrado, ou quando se percebe que o erro não está convergindo para o valor desejado e, portanto, a RNP deve ser remodelada.

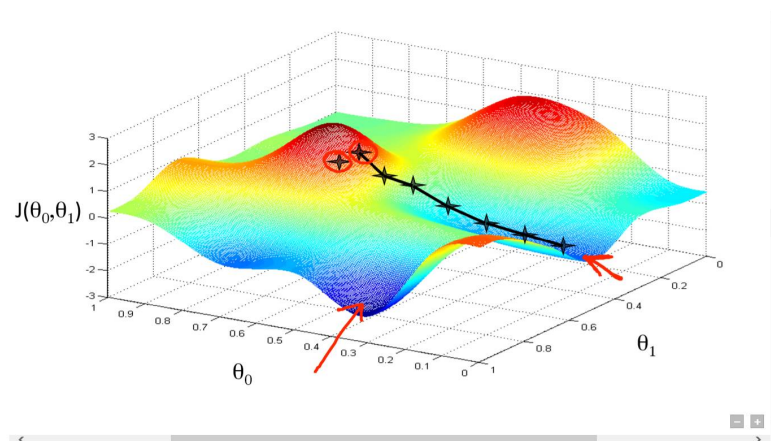


Figura 3- Exemplo de como o Método do Gradiente funciona. O "caminho" seguido pelo método encontrou um ponto de mínimo local, embora exista outro mínimo global (observar as indicações)

2.4.2 ÁRVORES DE DECISÃO

As respostas de um problema normalmente são influenciadas por diversos parâmetros de entrada, sendo que estes parâmetros possuem diversos graus de influência no resultado. O algoritmo de árvore de decisão executa sucessivas divisões nestes parâmetros, usando um critério bem definido para decidir como será esta divisão, formando assim classes de parâmetros. Cada divisão (decisão) deste algoritmo pode gerar uma folha, em que não é possível continuar com as subdivisões dos parâmetros em classes (é a resposta), ou pode gerar um ramo, em que as subdivisões podem prosseguir. Este processo recursivo se completa quando se obtém apenas folhas.

Os dados deste problema podem ser matematicamente representados como:

$$(x, Y) = (x_1, x_2, x_3, \dots, x_k, Y)$$

A variável dependente Y é o resultado que se deseja compreender ou prever, e o vetor x é composto por parâmetros de entrada que possuem diferentes graus de influência nesse resultado. Basicamente, existem dois tipos de árvores de decisão:

- Árvores de Classificação: o resultado a ser obtido é um valor discreto ou uma classificação
- Árvores de Regressão: o resultado a ser obtido é um valor numérico.

Por exemplo, para o problema de predição dos preços de uma residência, são muitos os parâmetros que têm influência no seu preço de venda. Considerando apenas dois parâmetros, número de quartos e metragem, a árvore de decisão poderia ser:

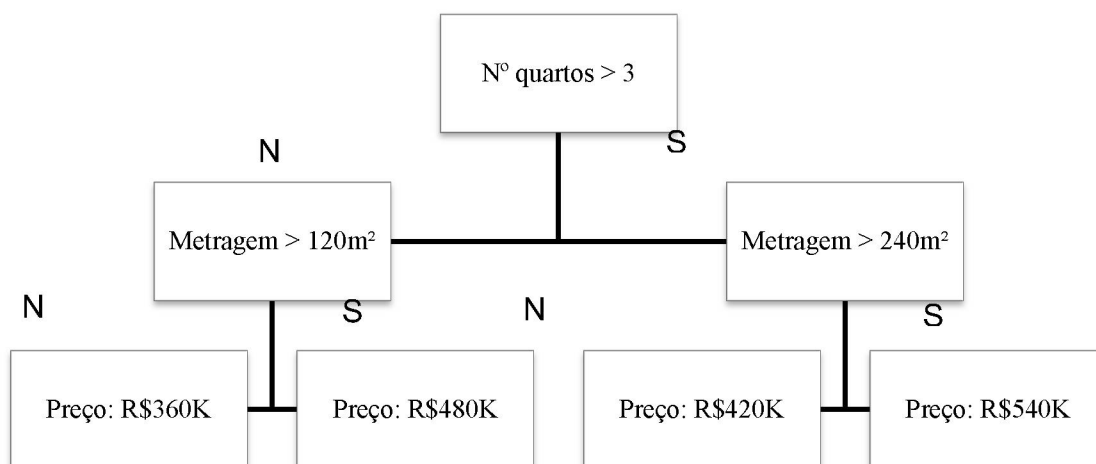


Figura 4 - Exemplo de uma Árvore de Decisão

É tarefa do algoritmo de geração da árvore de decisão construir esta hierarquia de decisões para se obter as respostas para o problema em questão. Isso acontece durante a fase de treinamento. Ou seja, assim como as RNAs, as árvores de decisão necessitam ser treinadas para poderem ser úteis em problemas de aprendizado de máquina.

A fase de treinamento apresenta algumas complexidades, dentre elas a definição do critério para proceder com as divisões dos parâmetros em grupos, o número de níveis a serem usadas na árvore e como descartar os parâmetros irrelevantes. A modelagem da árvore de decisão consiste justamente em transpor estas complexidades, visando a obtenção de um modelo genérico capaz de prever com precisão suficiente a resposta que se deseja.

2.4.3 MÁQUINAS DE VETORES DE SUPORTE

Os parâmetros que influenciam um determinado problema podem ser representados matematicamente através de um espaço de N dimensões, em que N representa a quantidade de parâmetros. Estes parâmetros, isoladamente ou através da iteração entre eles, acabam por resultar em diferentes valores para o problema em questão. O algoritmo denominado Máquina de Vetores de Suporte (MVS) propõe encontrar uma fronteira neste espaço em que os resultados são agrupados da forma mais otimizada possível. A representação gráfica em de um problema definido em um espaço bidimensional seria:

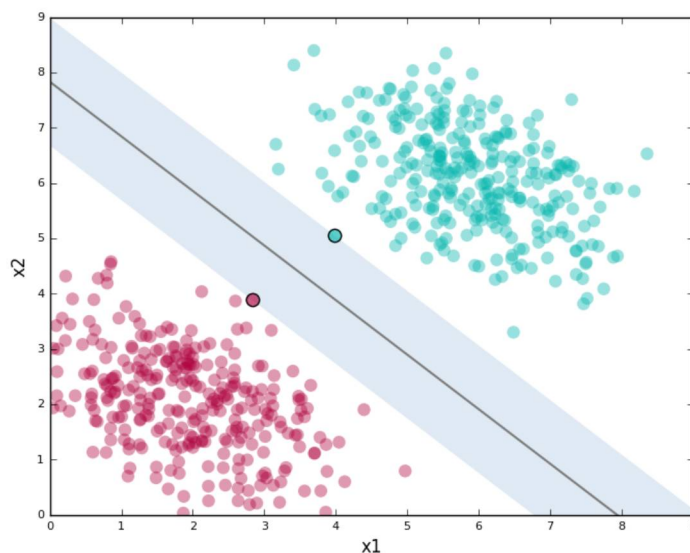


Figura 5 - Exemplo de como a Máquina de Vetores de Suporte funciona

Na figura acima, x_1 poderia ser o tamanho de um tumor e x_2 poderia ser a idade do paciente. Os pontos azuis indicariam um diagnóstico de tumor maligno. Por outro lado, os círculos na cor vermelha indicariam tumores benignos. O desafio do algoritmo é encontrar esta fronteira, de tal sorte a distância entre os pontos dos diferentes grupos, denominada margem, seja a maior possível. Em duas dimensões esta fronteira, tecnicamente chamada hiperplano, é uma linha. Eventualmente, pode ser impossível estabelecer um hiperplano perfeito entre os grupos, de modo que pontos vermelhos e azuis estejam do lado oposto da linha do hiperplano. Esta situação pode acontecer por diversas razões, as quais pode-se citar a presença dados espúrios ou mesmo devido à natureza não linear dos dados. Neste caso, o algoritmo procura minimizar os erros na classificação ao mesmo tempo que tenta maximizar a margem.

Da mesma forma que os métodos anteriormente citados, o MVS necessita ser treinado e validado antes de poder ser usado para fazer predições.

3 CENÁRIO ATUAL

Para levantar dados para o presente estudo, foi visitada uma concessionária de veículos da região de metropolitana de Curitiba. Foi constatado que a gestão do estoque possui uma variedade de itens fora de controle, gerando custos financeiros altos e problemas com itens obsoletos. O desbalanceamento do saldo de estoque é mostrado claramente com itens possuindo período de cobertura acima de 1 ano ou mais, no qual o passo seguinte é levado a obsolescência e conseqüentemente prejuízo para a empresa.

O descontrole da gestão é causado basicamente pela falta de comunicação, falta de regras para definição da compra de materiais e pela falta de dados históricos para tomada de decisão no momento de fechar a ordem de compra para o abastecimento do estoque.

Mesmo contando com softwares que dão informações mais do que detalhadas sobre o armazenamento de produtos ou com todos os controles feitos minuciosamente, os erros podem persistir, isso porque o setor de compras nunca deve trabalhar de forma isolada.

Além disso, a falta de comunicação entre o departamento comercial, peças de reposição e de compras podem comprometer o equilíbrio do saldo de estoque de todo o processo. Saber a frequência das vendas de um determinado item, bem como uma possível sazonalidade na procura será essencial.

Um das regras que a empresa não deve deixar de lado como um indicador de compras é a demanda de peças, um dos parâmetros que definem a condição de compra dos materiais de estoque, conhecido como “giro de material”.

O giro de materiais pode ser definido como o tempo em que cada item permanece dentro do estoque até que seja necessário fazer sua reposição. Quando a frequência de entregas é alta — em decorrência do grande volume de saída —, dizemos que o item é de alto giro. Acompanhar esse índice é necessário para que se saiba o tempo correto de acionar o fornecedor, evitando que o produto falte. Por outro lado, os itens de baixo giro indicam que aqueles produtos são pouco vendidos, o que significa que o ideal é evitar comprá-los, fazer as aquisições com um espaço de tempo maior ou adquiri-los em menor quantidade.

Para ilustrar, considere a figura abaixo. Na concessionária avaliada foi realizado um levantamento considerando o consumo médio de itens nos últimos 12 meses e a quantidade de itens em estoque na data atual. A divisão desta quantidade pelo consumo resulta num valor que se denomina tempo de cobertura de estoque. Trata-se de uma informação que ajuda a mensurar a eficácia do estoque. Idealmente, quanto mais itens se concentrarem em períodos de 0 a 3 meses, significa que o estoque gira mais rapidamente. Itens concentrados nos períodos de 3 a 6 meses ou de 6 a 9 meses seriam aceitáveis, dependendo da estratégia de estoque determinada pelo gerente de estoque. Por outro lado, quanto mais itens estão classificados como acima de 12 meses, significa que estes itens não estão sendo girados no estoque e indicam uma maior tendência a se tornarem obsoletos. Adicionalmente, representam custos financeiros de espaço físico ocupado e custos de oportunidade. Esta quantidade de itens não é desejada de forma alguma em uma concessionária cujo objetivo é uma gestão enxuta de estoques.

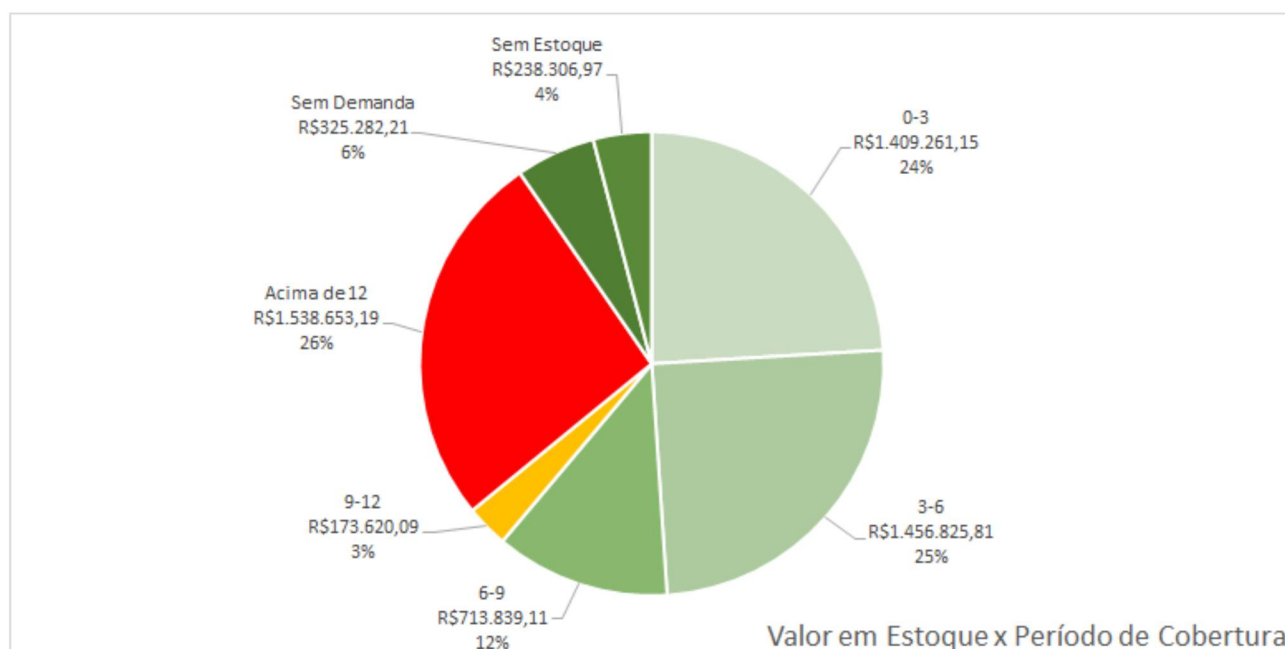


Figura 6 - Gráfico de Setores mostrando o perfil do estoque na Concessionária avaliada neste estudo. Nota-se que 26% dos itens estão acima de 12 meses do período de cobertura

Os dados históricos e o conhecimento de como extrair esses dados para obter informações de compra de materiais é um dos pontos chaves para evitar o chamado ruptura de estoque, que é a falta de mercadoria no momento que o cliente solicitou ou estoque em excesso. Outro fator muito comum é a falha na organização no estoque de materiais, tanto fisicamente quanto no gerenciamento pelo sistema de informações. Fisicamente, quando

podemos observar um SKU em diversos endereços de armazenamento, a falta de identificação e até mesmo a mistura de peças, podem gerar problemas de contagem no inventário e até a falta de espaço físico para outros materiais, além de potencialmente tornar os dados históricos suspeitos.

Na gestão das informações o problema é mais grave, pois além de depender das condições da organização física do setor de estoque, a falta de informação do produto ou a inserção errada de informação do material acaba gerando a ordem de compra de forma equivocada.

4 PROPOSTA DE SOLUÇÃO

Diante do que foi exposto, a gestão eficiente do estoque de peças de reposição em uma concessionária de veículos é um desafio imenso. Dentre as dificuldades encontradas, pode-se citar:

- Os veículos são compostos de muitas peças, podendo chegar a quase 30 mil peças, dependendo do modelo do veículo. Todas estas peças deveriam, em tese, estar prontamente disponíveis para utilização durante um procedimento de conserto.

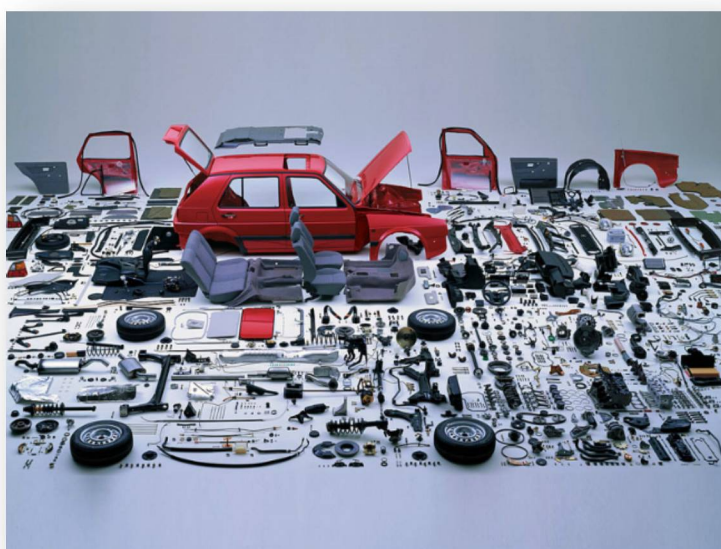


Figura 7 - Um carro pode ser composto de milhares de peças

- Algumas destas peças possuem características de demanda sazonal. Ou seja, possuem maior procura dependendo da estação do ano, pois sofrem influência do clima (temperatura, umidade, chuvas, poeira etc.)
- Não há fidelidade dos clientes para com as concessionárias. Ou seja, não há garantia do retorno do cliente procurando serviços de reparo para seus veículos. Estes clientes podem facilmente procurar a loja do concorrente para executar o procedimento de reparo. Tal comportamento prejudica uma previsibilidade da demanda futura de peças.
- A concessionária avaliada é composta de uma rede de diversas lojas, localizadas na capital e no interior do estado. Por política administrativa da rede, a unidade

matriz funciona como um centro de distribuição para as demais lojas. Assim, é esperado encontrar um desequilíbrio de estoque de peças, para o excesso, na unidade matriz, em relação às demais lojas.

- Grandes fornecedores de autopeças oferecem atraentes descontos para compras em volume, gerando um desequilíbrio no estoque de peças se a concessionária optar por beneficiar-se destes descontos.
- O tempo de entrega do pedido por parte dos fornecedores é bastante aleatório, podendo variar de alguns poucos dias para semanas, dependendo do volume solicitado e do tipo/complexidade da peça.
- Fabricantes podem aprimorar ou alterar completamente as especificações das peças durante a vigência da comercialização dos veículos, potencialmente provocando o surgimento de peças obsoletas, caso estas já tenham sido compradas pelas concessionárias antes desta alteração.

Assim, considerando todas estas dificuldades, propõe-se desenvolver um algoritmo que utilizará a tecnologia de Aprendizado de Máquina, citada nas seções anteriores, para tentar prever a demanda futura de peças de reposição nos veículos comercializados pela concessionária.

Este algoritmo coletará dados de diversas fontes, sendo que estas divididas entre contextuais – relacionadas com o ambiente ou o contexto comercial histórico em que a concessionária está inserida, e intrínsecas – relacionadas com os diversos dados coletados por sensores instalados nos veículos.

Para os dados contextuais, tem-se:

- Dados históricos da comercialização de veículos praticada pela concessionária nos últimos 10 anos (considerando que a idade média da frota de veículos no Brasil é de 9,7 anos, segundo informações do Sindipeças – Sindicato Nacional da Indústria de Componentes para Veículos Automotores)
 - Serão coletados também dados históricos da comercialização de veículos das concessionárias concorrentes, mas do mesmo fabricante, na região de influência comercial da concessionária, e para o mesmo período temporal, obtidos através da FENABRAVE

- Dados históricos da venda de autopeças, obtidas através do SINDIPEÇAS
- Dados climáticos, tais como pressão atmosférica, temperatura média diária, umidade, precipitação pluviométrica, índice de poluição, qualidade do ar, obtidas através do SIMEPAR ou IAP.
- Adicionalmente, pode-se também utilizar como dados contextuais o Índice de Manutenção Veicular divulgados pelo grupo CESVI – Centro de Experimentação e Segurança Viária.

Para os dados intrínsecos aos veículos:

- Dados referentes à dinâmica veicular (velocidade média, velocidade instantânea, aceleração média, aceleração instantânea)
- Dados referentes à topografia de circulação do veículo (aclives, declives, planos)
- Dados referentes à parte elétrica/eletrônica do veículo (carga da bateria, nº de ciclos partidas da bateria por dia, registro instantâneo do consumo de eletricidade pelos diversos componentes)
- Dados referentes ao motor a combustão (nº de horas em uso do motor, curva de RPM, curva de torque, curva de temperatura do líquido de arrefecimento, curva da temperatura do óleo de cárter)
- Dados referentes aos gases de combustão, coletados através da sonda lambda (esta sonda mede a eficiência da queima do combustível)

Tais dados intrínsecos dos veículos serão coletados por intermédio de um dispositivo denominado OBDII, a ser instalado em uma tomada específica e que todos os veículos equipados com injeção eletrônica possuem. Este dispositivo normalmente é equipado com um transmissor *bluetooth*, que pode enviar os dados coletados para qualquer outro dispositivo compatível com a tecnologia *bluetooth*, tais como *scanners* de oficinas, smartphones e *laptops*. Na presente proposta de solução, estes dados seriam coletados através de um aplicativo de smartphone, que então enviaria os dados para um banco de dados “em nuvem”, para posterior análise pelo algoritmo de Aprendizado de Máquina.

Finalmente, após executar a análise de todos estes dados (contextuais e intrínsecos) pelo algoritmo, os resultados serão apresentados num *dashboard*, com as

principais métricas coletadas pelo sistema e com recomendações para níveis de estoque de peças otimizados segundo os critérios e objetivos desejados pela concessionária.

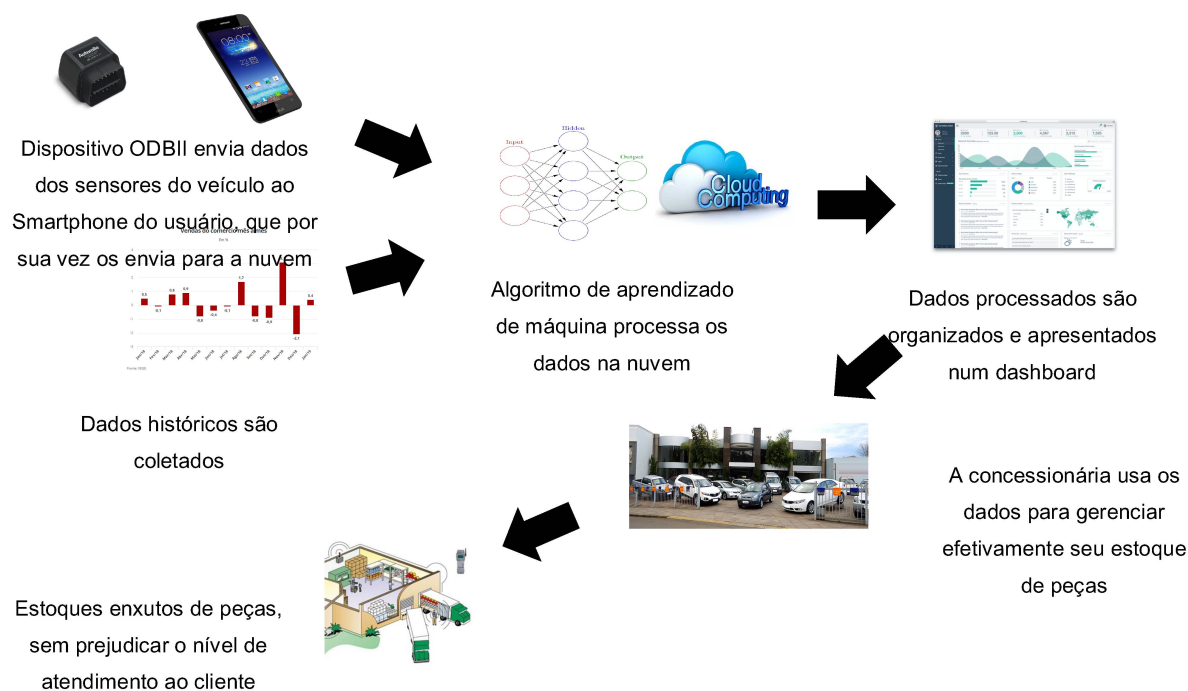


Figura 8 - Esquema mostrando como funcionaria a proposta de solução para o problema identificado na concessionária avaliada

4.1 IMPLEMENTAÇÃO

O primeiro passo para a implementação desta proposta de solução consiste em desenvolver a tecnologia necessária para coletar os dados de entrada do algoritmo de aprendizado de máquina. Tem-se basicamente duas fontes de dados a serem coletadas: os dados contextuais e os dados intrínsecos.

Conforme comentado, os dados contextuais são relativos ao ambiente ou contexto ao qual o veículo está circulando. São importantes, pois é sabido que condições climáticas, como umidade ou presença de maresia, tem influência decisiva na durabilidade de algumas peças dos veículos. Para coletar os dados contextuais, será desenvolvido um aplicativo de celular que irá registrar a data e a localização por GPS da circulação do veículo. Desta forma, pode-se cruzar estes dados com os registros históricos do clima para determinar as condições específicas na localização de circulação do veículo e na data desta circulação.

Um exemplo de tabela de dados de localização que seriam coletados pelo aplicativo, com o veículo em funcionamento, pode ser visto abaixo:

```
data;hora;longitude;latitude
20191002;103000;-25.423236;-49.304937
20191002;104500;-25.498656;-49.362598
20191002;110000;-25.563214;-49.395624
20191002;111500;-25.593267;-49.415275
20191002;113000;-25.623366;-49.452369
...
```

O aplicativo armazenaria estes dados e, uma vez por dia, faria o upload para um sistema a ser desenvolvido “na nuvem” para consolidar estes dados e cruzar com os dados climáticos (a serem coletados a partir de entidades governamentais de monitoramento do clima, como o SIMEPAR ou IAP, dentre outras), resultando numa outra tabela, desta vez pronta para ser analisada pelo algoritmo de aprendizado de máquina

```
data;hora;temperatura;umidade;pressão;precipitação;maresia;dióxido_enxofre;dióxido_nitrogenio
20191002;103000;25;0.45;1012;0.0;0;0;0
20191002;104500;26;0.46;1015;0.0;0;0;0
20191002;110000;26;0.46;1014;0.0;0;0;0
20191002;111500;27;0.45;1014;0.0;0;0;0
20191002;113000;27;0.45;1013;0.0;0;0;0
...
```

Figura 9- Tabela com um exemplo de como seriam os dados do ambiente no qual o veículo pode estar circulando

Assim, o algoritmo pode avaliar os impactos destes parâmetros na vida útil dos componentes do veículo. Obviamente, haverá de ser necessário alimentar o algoritmo com uma quantidade suficiente de dados, com seus respectivos impactos na durabilidade dos componentes dos veículos para que o algoritmo possa “aprender” estes impactos.

Em relação aos dados intrínsecos, estes seriam coletados através da leitura das informações provenientes da porta OBDII dos veículos, que em seguida seriam transmitidos via *Bluetooth* pelo aplicativo de celular do usuário. Da mesma forma, periodicamente, tais dados seriam enviados para o sistema “na nuvem”, para em seguida serem consolidados.

Figura 10 - Tabela com um exemplo de como seriam os dados coletados através dos sensores dos veículos

A figura abaixo mostra um exemplo do aspecto da tabela de informações lidas da porta OBDII

```
data;hora;temp;rpm;acel;lambda;tensão;temp_arref
20191002;103000;71;3500;0.7;11.7;45
20191002;104500;77;3600;0.8;12.0;55
20191002;110000;80;3700;0.9;11.1;56
20191002;111500;81;2000;0.7;11.2;49
20191002;113000;90;2900;0.9;11.9;48
...
```

Alguns destes dados, tais como velocidade, aceleração e altitude, seriam obtidos diretamente dos sensores localizados (GPS, acelerômetro) no aparelho smartphone do usuário.

Por tentar exemplificar como funcionaria na prática toda essa coleta de dados, considere a ilustração esquemática abaixo mostrando a região de circulação de dois veículos idênticos (fabricante, modelo), adquiridos na mesma data, mas que circulam em diferentes regiões e são de proprietários diferentes:

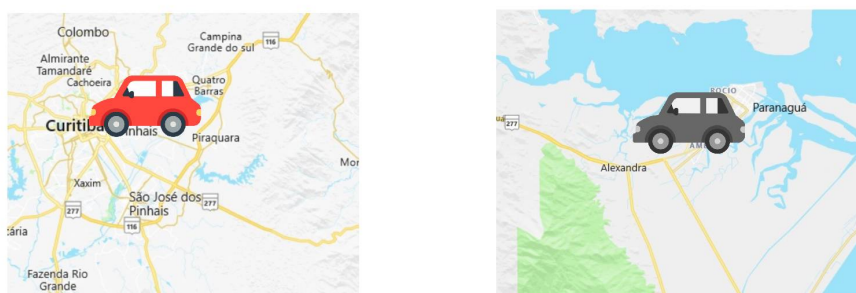


Figura 11 - Exemplo ilustrando a influência da região de circulação nas peças de um veículo

Os dados registrados para o veículo vermelho indicam que 90% do tempo o histórico de localização aponta circulação restrita a região metropolitana de Curitiba, com pouca variação de topografia, sujeita às condições climáticas e de tráfego locais. Já para o veículo cinza escuro, os dados registram grande circulação em estrada, entre Curitiba e o litoral paranaense, com pernoite no litoral. Portanto esse veículo experimenta grande variação de topografia (900+ metros), além de velocidade média bastante elevada. Com base nestas informações, o algoritmo pode inferir que os desgastes dos componentes dos

dois veículos serão bastante diferentes. O veículo vermelho, apesar de apresentar um hodômetro menor, seus dados revelam uma maior intensidade de ciclos de aceleração e desaceleração. Esse fato está relacionado com o hábito de guiar do condutor, mas também tem a ver com o fato da região de circulação ser predominantemente urbana. Já o veículo cinza escuro apresenta menos ciclos de aceleração e desaceleração, mas por outro lado estas são potencialmente mais intensas, devido às características do tráfego em estrada.

Quais seriam os efeitos destas condições na pastilha de freio destes veículos? O algoritmo de aprendizado de máquina pode indicar um desgaste mais intenso no veículo vermelho, por sofrer com a maior frequência dos ciclos de aceleração/desaceleração. Ao comparar com os dados de veículo similar, cuja condição de circulação é também similar, o algoritmo indicaria a necessidade de uma troca em dois meses, mantidas as condições atuais de circulação. Tal informação seria mostrada no *dashboard* da concessionária, que preventivamente poderia acrescentar uma unidade de um jogo de pastilhas de freio para o modelo do veículo em questão na programação futura de compras junto ao fornecedor de peças.

4.2 PÚBLICO ALVO E FORMA DE ENTREGA DO PRODUTO

Conforme mencionado na seção “Introdução”, o público alvo para a presente proposta de solução são concessionárias de veículos automotores que experimentam dificuldades em controlar de forma enxuta seu estoque de autopeças de reposição, usadas nos reparos dos veículos.

Os resultados das análises do algoritmo de aprendizado de máquina serão apresentados na forma de um *dashboard*, mostrando todas as principais métricas necessárias para a sua correta interpretação das condições atuais do estoque de autopeças. O gestor de estoque poderá observar também possíveis cenários de demanda por autopeças num futuro próximo, considerando probabilidades calculadas pelo algoritmo. Assim, conhecendo o estoque atual e seu provável comportamento futuro, o gestor poderá tomar as decisões mais assertivas possíveis para manter os estoques enxutos, porém sem prejudicar o nível de serviço ao cliente.

Estas concessionárias deverão contratar o serviço de implantação do produto, que será entregue da seguinte forma:

- Sistema (*web application*) que contém o algoritmo de aprendizado de máquina e que permite sua parametrização, através de uma interface gráfica, segundo as necessidades do cliente
- Implantação deste sistema como um serviço de internet acessível através de provedores de computação em nuvem, como a Amazon AWS ou Microsoft Azure. Este provedor irá fornecer toda a infraestrutura (servidores, banco de dados etc.) necessários para este sistema
- Treinamento para interpretação das métricas apresentadas no dashboard.
- Suporte pós-venda
- Dispositivo OBDII, a ser instalado em cada veículo comercializado pela concessionária.
- Aplicativo de smartphone que, além de coletar os sensores próprios do aparelho, coletará também os dados lidos pelo dispositivo OBDII e os transmitirá para o sistema na nuvem.

Além das concessionárias de veículos, outros possíveis interessados nesse sistema de gestão de estoque auxiliado por aprendizado de máquina seriam grandes frotistas de veículos de empresas de serviço público (companhias de água e saneamento, companhias de luz etc.) que possuem oficinas próprias para reparos. Grandes empresas de aluguéis de veículos também poderiam compor um potencial público alvo.



Figura 12- Potenciais clientes da solução a ser desenvolvida

4.3 INVESTIMENTO NECESSÁRIO E ESTRATÉGIA DE COMERCIALIZAÇÃO

Esta proposta de solução está fortemente baseada em tecnologias como aprendizado de máquina, computação em nuvem, desenvolvimento de códigos para dispositivos móveis etc. Portanto, boa parte do investimento necessário está relacionado com infraestrutura de tecnologia de informação. Grandes provedores de serviços de TI em nuvem, como a Amazon AWS ou Microsoft Azure oferecem grande parte da tecnologia necessária a um custo acessível. O principal benefício de se usar os serviços destes provedores está na escalabilidade: sempre que necessário, pode-se expandir ou retrain a capacidade para se adequar as necessidades dos clientes. Outro benefício está na variedade de ofertas de serviços destes provedores. Além de ofertar a capacidade de computação sob demanda, oferecem também os serviços de bancos de dados escaláveis e garantem, por contrato de nível de serviço, a alta disponibilidade destes serviços. Ou seja, dificilmente o sistema a ser desenvolvido ficará indisponível por falha do servidor localizado na nuvem.

Deverá ser contratada uma equipe de desenvolvedores com experiência em computação em nuvem e banco de dados para escrever os códigos requeridos para a implantação do sistema. O esforço necessário para esta equipe estará dividido nas seguintes atividades:

- Desenvolvimento do algoritmo de aprendizado de máquina;
- Desenvolvimento do Banco de Dados que irá armazenar os dados coletados;
- Desenvolvimento de um aplicativo para smartphone que coletará os dados dos sensores dos veículos, através do dispositivo OBDII, bem como do próprio smartphone;

Num primeiro momento, o dispositivo OBDII a ser instalado nos veículos não necessitará de nenhuma customização, podendo ser adquirido diretamente dos fornecedores de produtos eletrônicos. Posteriormente, uma boa ideia seria desenvolver um

ODBII customizado capaz de coletar, armazenar e transmitir os dados sem a necessidade de se utilizar o aparelho smartphone do usuário.

Uma característica importante desta proposta de solução está em seu caráter de consultoria e parceria. Visto que o algoritmo de aprendizado de máquina necessita ser treinado por algum tempo para se começar a obter resultados confiáveis, a concessionária contratante precisará acreditar na tecnologia antes de começar a colher os frutos. Trata-se, portanto, de um desenvolvimento conjunto, que envolve um período de coleta de dados, seguido por um período de ajuste e treinamento usando resultados já conhecidos para finalmente beneficiar-se das capacidades preditivas do algoritmo já no período final de implantação. E mesmo após a implantação, visto o caráter dinâmico do mercado de autopeças para veículos, eventuais ajustes ainda se farão necessários.

Esta característica tem impacto na estratégia de comercialização do produto desta proposta. Simplesmente entregar ao cliente o produto e abandonar ao cliente não funcionaria. Assim, a entrega da solução deverá ser composta de uma assinatura de um contrato de serviço, em que o cliente faria desembolsos mensais para beneficiar-se da tecnologia. Este contrato deverá prever algumas fases de implantação, considerando as características citadas anteriormente:

- Fase I: Estudo inicial, coleta das informações do ambiente e do mercado em que a concessionária está inserida. Deve-se coletar também as expectativas, metas e objetivos da concessionária para o uso da ferramenta
- Fase II: Implantação da solução preliminar ajustada para as necessidades manifestadas na fase anterior. É nesta fase deverá ser executada o treinamento do algoritmo de aprendizado de máquina. O período requerido para esta fase dependerá da qualidade e quantidade dos dados coletados e da assertividade dos resultados do algoritmo.
- Fase III: Aprimoramento da solução. Uma vez que o algoritmo está funcionando satisfatoriamente, esta fase na realidade é o pós-venda, e compreende a manutenção do sistema, possíveis ajustes e aperfeiçoamentos adicionais, bem como suporte e treinamento dos funcionários da concessionária que irão operar o sistema.

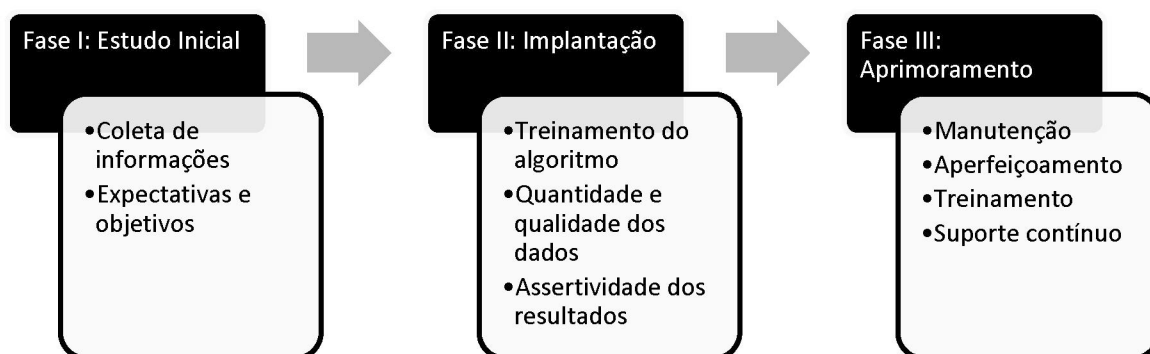


Figura 13 - Fases de implantação da solução

Os investimentos necessários para iniciar a empresa que irá desenvolver a solução do algoritmo de aprendizado de máquina incluem os custos de toda a infraestrutura de aluguel de poder computacional em nuvem, de uma equipe de desenvolvedores, e de um espaço em um escritório compartilhado. Além disso, deve-se considerar um período inicial em que a equipe deverá escrever o código base para fundamentar todo o desenvolvimento posterior. O tempo necessário é difícil de estimar, pois depende das habilidades do time de desenvolvedores a ser contratado, mas estima-se que seriam necessários ao menos seis meses para se ter um protótipo funcional capaz de ser apresentado à uma possível primeira concessionária cliente.

Esta primeira concessionária cliente, na realidade, terá mais o papel de uma patrocinadora da solução. Ela permitirá colocar em prática o algoritmo, definindo os objetivos e expectativas, e cedendo os dados reais necessários para o treinamento da primeira versão do algoritmo. Portanto, em troca deste patrocínio, receberá todo esse desenvolvimento preliminar da solução sem custo algum, em troca destes dados iniciais reais do mercado e da oportunidade de se colocar em prática o algoritmo. Novamente, o tempo requerido para este período é difícil de estimar, mas imagina-se que 12 meses seriam suficientes para permitir uma boa coleta de dados, garantir que esta coleta de dados está usando informações úteis e relevantes de entrada para o algoritmo, e para permitir também os ajustes que se fizerem necessários.

Em resumo, seriam 6 meses de trabalho interno somados com 12 meses de uma implantação preliminar numa concessionária patrocinadora, resultando em 18 meses. Ao se adicionar mais 6 meses de segurança, seriam 24 meses de desenvolvimento inicial, validados por esta concessionária patrocinadora.

Com relação aos números do investimento financeiro para esse período de incubação inicial da solução, estima-se que seriam necessários cerca de R\$350 mil, referentes à contratação de uma equipe de 2 desenvolvedores de código, ao aluguel da infraestrutura de computação em nuvem e ao aluguel de um espaço em um escritório compartilhado.

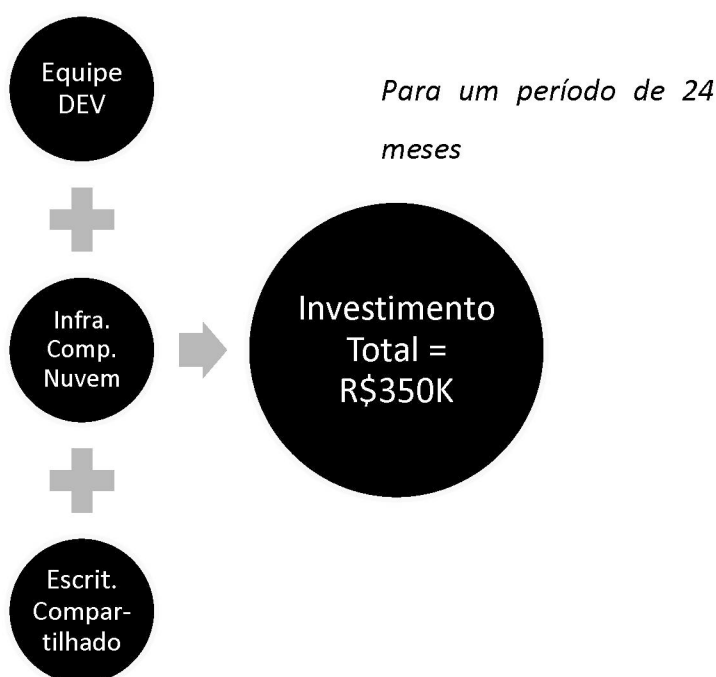


Figura 14 - Investimento necessário para implantar a solução, considerando um período de 24 meses

Após esse período previsto inicial, já quando a solução estiver validada, amadurecida e pronta para ser oferecida para outros clientes, a estratégia de comercialização e monetização, conforme já comentado, será via uma assinatura mensal de serviços. Ou seja, o cliente adquire uma assinatura que contém todos os serviços necessários para a implantação e manutenção da solução, e que contém o sistema com a solução do algoritmo de aprendizado de máquina. Estima-se que esta assinatura custe ao cliente cerca de R\$1,5 mil mensais, o que é um custo bastante competitivo ao se considerar os potenciais benefícios ao longo do tempo. Com este valor, a empresa a ser criada teria

de possuir em sua carteira de clientes no mínimo 10 contratos de serviços fechados para obter o ponto de equilíbrio, pois estima-se que os custos mensais para manter a empresa seriam de R\$15 mil.

Finalmente, em termos de retorno esperado sobre o investimento, considerando que o montante em estoque na concessionária avaliada nesse estudo é de cerca de R\$500 mil, se ao longo de 12 meses houver uma economia de apenas 10% desse valor, ou seja, R\$50 mil, já seriam suficientes para cobrir os custos anuais da assinatura da solução e ainda sobraria um troco:

$$R\$1.500 \times 12 \text{ meses} = R\$18.000 \rightarrow \text{custo anual da solução}$$

$$R\$50.000 - R\$18.000 = R\$32.000 \rightarrow \text{valor economizado anualmente}$$

$$\frac{R\$32.000}{R\$18.000} \times 100 = 177.7 \% \rightarrow ROI$$

Vale mencionar que estas estimativas estão extremamente conservadoras, tanto com relação ao custo da assinatura mensal, visto que não há soluções semelhantes disponíveis no mercado, quanto com respeito à expectativa de economia. Aliás, estes custos ainda não incluem valores relacionados à economia de espaço físico necessário para alocar os itens de estoque, custos de depreciação, custos de mão de obra para manipulação/seleção/conferência e os custos de oportunidade financeira. Caso seja aplicado um maior rigor nestas contas, certamente o valor economizado anualmente será bem mais atraente.

5 CONCLUSÃO

A tecnologia de aprendizado de máquina vem sendo adotada cada vez mais por diferentes áreas do conhecimento e setores econômicos, variando desde climatologia, em que é empregada no auxílio da previsão de tempo em uma área específica com o objetivo de reduzir os riscos de uma cultura agrícola, até na segurança pública, em que análise de imagens vem sendo usada para identificar padrões de comportamento suspeitos e assim antecipar uma ação por parte das autoridades que pode prevenir uma ocorrência policial ou mesmo para evitar vítimas.

A razão deste crescimento está, obviamente, no aumento do poder computacional dos computadores modernos, principalmente com o advento do processamento através das unidades GPU – *Graphical Processing Unit* – que são compostas por milhares de processadores de fluxo – *Stream Processors* – que permitem o processamento paralelo das matrizes numéricas usadas pelos algoritmos de aprendizado de máquina. Por exemplo, uma placa GPU NVIDIA GeForce RTX 2080 Ti possui cerca de 4300 cores – ou processadores – capazes de atingir a velocidade de cálculo de 13,45 TFLOPS para números de ponto flutuante de 32 bits. É bastante poder computacional para uma única placa. Além disso, pode-se conceber uma arquitetura de hardware em que diversas destas placas são utilizadas em conjunto, formando um cluster computacional com enorme velocidade e capacidade de processar dados.



Figura 15 - Placa gráfica GPU usada em algoritmos de aprendizado de máquina

Outro desenvolvimento recente que está favorecendo a popularização dos algoritmos de aprendizado de máquina se refere ao imenso aprimoramento das bibliotecas de código que foram desenvolvidas nos últimos anos. Conforme citado na introdução, a tecnologia “aprendizado de máquina” não é recente, existe há várias décadas. Mas até recentemente, no entanto, manteve-se circunscrita ao ambiente acadêmico. Foram iniciativas de grandes empresas de tecnologia, como a Google e Facebook, que disponibilizaram gratuitamente ao público em geral o acesso às avançadas bibliotecas de código com diversos algoritmos já implementados, além de disponibilizar extensa documentação sobre como utilizar.

Finalmente, outro aspecto da tecnologia que irá impulsionar de uma vez por todas a adoção da tecnologia de aprendizado de máquina está do advento de serviços de computação em nuvem, como as ofertadas pela Amazon AWS ou Microsoft Azure. A placa GPU da NVIDIA citada anteriormente custa cerca de R\$8.000. Um computador completo, que utiliza esta placa deverá custar em torno de R\$15.000. Não é um valor baixo. Por uma fração deste valor, pode-se ter acesso a uma máquina virtual disponibilizada na nuvem com um poder computacional muitas vezes superior ao computador mencionado. O segredo desta aparente contradição reside no fato de que ao se utilizar os serviços da nuvem, paga-se apenas pelo tempo de processamento dispendido pelo do algoritmo de aprendizado de máquina, reduzindo desta forma os custos por resultado obtido.

Tem-se, portanto, um cenário bastante favorável ao empreendedorismo usando a tecnologia de aprendizado de máquina. Talvez o único ponto desfavorável neste “oceano azul” seria a escassez de mão de obra especializada. O cientista de dados, nome da profissão responsável pela modelagem, treinamento, teste e aperfeiçoamento de algoritmos de inteligência artificial, ainda é um profissional raro de se encontrar no mercado de trabalho. Embora os algoritmos estejam prontamente disponíveis e, em teoria, sejam relativamente fáceis de se utilizar por um desenvolvedor comum, o grande segredo desta profissão está na modelagem e escolha mais adequada do algoritmo para o problema em questão, bem como na habilidade de se interpretar os resultados e implementar os ajustes necessários para assegurar a máxima eficiência do código. A eficiência é importante, pois quanto mais dados, naturalmente mais precisos e confiáveis serão os resultados. E pode-se aumentar a eficiência alugando-se mais poder computacional e/ou contratando cientistas de dados habilitados.

No presente objeto de estudo deste artigo, é nítido que esta tecnologia seria extremamente benéfica para promover uma gestão de estoque de peças eficiente em uma concessionária de veículos. Embora muitos dados sejam coletados diariamente por diversos sistemas, tanto os relativos aos softwares de gestão como CRM ou ERP, como aqueles vindos dos próprios veículos, estes dados normalmente não são utilizados com o objetivo de promover uma gestão de estoques eficaz. Acaba-se por negligenciar a relevância destes dados, a princípio não apenas por incompetência da gestão da concessionária, mas principalmente pelo desconhecimento de que a coleta, modelagem, correlação e interpretação dos diversos dados que acabam circulando pelas atividades do negócio, muitas vezes soltos e despercebidos, podem se tornar fundamentais para se descobrir padrões que, por seu turno, são fundamentais para a vantagem competitiva da empresa perante os concorrentes.

Iniciativas de utilização da tecnologia de aprendizado de máquina para se prever o comportamento futuro de determinada variável tem sido usada com sucesso nos últimos anos. Um exemplo clássico está na área de manutenção de máquinas: usam-se diversos sensores para coletar dados que, em última análise, são utilizados para se determinar uma agenda de manutenção que por fim reduz drasticamente os custos associados. O conceito apresentado neste artigo, fazendo-se uma extrapolação para um conjunto de veículos, utiliza uma ideia semelhante: prever com algum grau de confiabilidade quando um componente do veículo irá falhar e tomar as ações que se fizerem necessárias (principalmente relacionadas com o estoque). Ainda não foi mencionado aqui, mas para o condutor do veículo também será extremamente vantajoso, pois ele terá um plano de manutenção dinâmico e personalizado à sua forma de dirigir, o que ao fim acaba reduzindo seus custos de com o veículo.

Uma ressalva importante e necessária vale ser mencionada. Embora pareça muito atraente e infalível usar de um algoritmo de aprendizado de máquina para auxiliar na otimização da gestão de estoque, não se pode cometer o erro de se otimizar o caos. É imprescindível organizar e sanitizar os bancos de dados da concessionária (CRM, ERP, estoque) antes de se tomar qualquer ação na direção de usar tecnologias auxiliadas por computador. Afinal, caso sejam usados dados inválidos, os resultados impreterivelmente serão inválidos também. A seção Cenário Atual revelou que para a concessionária avaliada neste artigo este passo preliminar seria imprescindível. E certamente não é a única

empresa com este tipo de dificuldade em se manter os dados corretos e validados. Desta forma, antes de ser um problema, ter-se-ia uma oportunidade de oferecer um serviço que faria parte da proposta comercial a ser oferecida para os potenciais clientes. Estes, por sua vez, apenas por implantar uma gestão enxuta tradicional, mesmo que ainda não se sintam confortáveis em utilizar o auxílio de decisões recomendadas por computador, já se beneficiariam enormemente.

Vale ressaltar a confiança dos autores deste trabalho no potencial benéfico tecnologia de aprendizado de máquina, para as mais diferentes áreas e setores da economia. Certamente é uma tecnologia que libertará a humanidade de tarefas repetitivas e tediosas, permitindo o uso melhor do tempo em atividades criativas e artísticas. Quando associada com a tecnologias como robótica, genética e nanotecnologia, tem-se como resultado um cenário em que, eventualmente em algum ponto no futuro, a sociedade como conhecemos atualmente sofrerá um impacto profundo. É o que os futurólogos chamam de ponto de singularidade. Os pessimistas dizem que a partir deste ponto a humanidade será irrelevante. Os otimistas dizem que nesse ponto a humanidade atingirá um grau de desenvolvimento nunca experimentado anteriormente e que não haverá conflitos, visto que a abundância de recursos será tamanha que não justificará qualquer desentendimento.

Como conclusão final, e voltando à realidade do presente trabalho, os autores acreditam também que a tecnologia de aprendizado de máquina não é uma questão de se vai adotar, mais sim de quando vai adotar. Foram citados neste texto diversos exemplos em alguns setores, mas o fato é trata-se de uma tecnologia genérica e pode ser implantada em praticamente qualquer área. Restam aos gestores a decisão da adoção ou não desta tecnologia antes de seus concorrentes.

REFERÊNCIAS

Bishop, C. M. (2006), **Pattern Recognition and Machine Learning**, Springer

De Mauro, Andrea; Greco, Marco; Grimaldi, Michele (2016). **A Formal definition of Big Data based on its essential Features**. Library Review. 65 (3): 122–135.
doi:10.1108/LR-06-2015-0061

Dhairya Kumar - **Demystifying Support Vector Machines** - Towards Data Science – Acesso em agosto de 2019 - <https://towardsdatascience.com/demystifying-support-vector-machines-8453b39f7368>

Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020. Gartner. 12 de dezembro de 2013

IAP – **Dados Ambientais** - <http://www.iap.pr.gov.br/pagina-59.html> - Acesso em setembro de 2019

Neural Network Using Python and Numpy https://www.python-course.eu/neural_networks_with_python_numpy.php - Acesso em agosto de 2019

Portal iCarros – **Idade da Frota de Veículos no Brasil** – Acesso em setembro de 2019 - <https://www.icarros.com.br/noticias/geral/frota-de-carros-no-brasil-e-a-mais-velha-em-18-anos/26397.html>

Russell, Stuart; Norvig, Peter (2003 [1995]). **Artificial Intelligence: A Modern Approach** 2 ed.: Prentice Hall.

SINDIPEÇAS - <https://www.sindipecas.org.br/home/> - Acesso em agosto de 2019

SIMEPAR – **Dados climáticos** – Acesso em setembro de 2019 - http://www.simepar.br/prognozweb/simepar/dados_estacoes/25264916

60 Must-Know Cybersecurity Statistics for 2019 - <https://www.varonis.com/blog/cybersecurity-statistics/> - Acesso em agosto de 2019