

# SVR para Predição de Ações do Facebook.

Guilherme Locca Salomão<sup>1</sup>, João Victor Mendes Freire<sup>1</sup>, Luís Felipe Corrêa Ortolan<sup>1</sup>

<sup>1</sup>Departamento de Computação - Universidade Federal de São Carlos (UFSCar)  
São Carlos - SP - Brazil

guilssalomao@gmail.com, joaovictormfreire@me.com, luisfelipe.ortolan@gmail.com

**Abstract.** *With so many variables and the price varying day by day, predicting value stocks can be a hard task. Many algorithms propose to predict the value of stocks using different methods. In this article we will look at a simple prediction method using Support Vector Regression (SVR). To predict new data, the algorithm attempts to approximate a function that describes well the previous data, therefore possibly describing the next points well.*

**Resumo.** *Prever o valor de ações pode ser uma tarefa difícil com tantas variáveis e o preço variando dia após dia. Muitos algoritmos se propõem a prever o valor das ações usando os mais diversos métodos. Neste artigo veremos um método de previsão simples usando Support Vector Regression (SVR). Para prever novos dados, o algoritmo tenta aproximar uma função que descreve bem os dados anteriores, assim, possivelmente irá descrever bem os próximos pontos.*

## 1. Apresentação do Problema

O mercado de ações é um ambiente onde muito dinheiro é movimentado. Investidores buscam investir em ações que aumentem seu preço e aumentem o mais rápido possível. Entretanto, nem todas elas sempre sobem, podendo desvalorizar bastante, e acabar gerando perdas aos seus investidores.

Para um investidor, o ideal seria saber quando uma ação valorizará seu preço e quando desvalorizará. Existem diversas variáveis que influenciam o valor das ações, como o momento econômico do país sede da companhia, ou se tal empresa está passando por um período de lucros ou perdas. Com tantos fatores que podem influenciar o mercado, é praticamente impossível que um acionista saiba exatamente quando uma ação será valorizada ou não.

Muitos algoritmos tentam usar do Aprendizado de Máquina para auxiliar investidores na previsão das variações dos preços das ações e assim saber quando comprá-las ou vendê-las. Existem diversos algoritmos que utilizam informações diferentes na tentativa de prever de forma mais precisa o valor de um título no mercado de ações. Nesse projeto, veremos um algoritmo de previsão de preços de ações bastante simples, que se utiliza apenas dos valores anteriores de uma determinada ação para tentar prever o valor desta no futuro.

## 2. Explicação Teórica

Como dito anteriormente, o algoritmo escolhido para apresentação nesse projeto é um algoritmo simples de predição dos valores da bolsa. Ele considera os valores que a ação

assumiu durante um período de um mês. A forma que o algoritmo busca fazer a previsão dos valores é primeiro tratando as informações como pontos no plano cartesiano.

Para tratar os dados das ações como parte do plano cartesiano, assume-se, nesse caso, que um dos eixos desse plano representa a data e o outro eixo representa o valor dessa ação. Interpretando as informações dessa maneira, obtém-se vários pontos, que podem ser utilizados para aproximar funções que representem como foi a oscilação do valor daquela ação durante o mês informado. Tendo uma curva que adaptou bem os pontos anteriores, é bem provável que ela venha a adaptar novos pontos, ou nesse caso valores, de maneira correta.

A técnica de Aprendizado de Máquina a ser utilizada é a Support Vector Regression (SVR). Essa técnica combina Support Vector Machine (SVM), uma técnica de Aprendizado de Máquina para classificação com Regressão Linear.

SVM são utilizadas para encontrar um hiperplano que separa duas classes em um gráfico. O objetivo é encontrar o hiperplano que maximiza a distância entre o plano e os elementos mais próximos de cada classe, chamados vetores. A técnica de Regressão Linear busca diminuir uma função de custo (utilizando algo como uma Descida de Encosta, ou Gradient Descent).

A técnica de SVR combina ambos os princípios, mas não para classificar grupos, e sim para obter o próximo valor real em dado conjunto. Procura-se minimizar o erro dentro de um intervalo procurando o hiperplano que contém o maior número de pontos e considerando apenas eles no modelo.

Nem todos os problemas, como é o caso da previsão de ações, são lineares. Assim sendo, utiliza-se funções para mapear os valores para um modelo linear. Tais funções são chamadas de kernel.

No algoritmo em questão, três tipos de função (kernel) foram usados para aproximar pontos anteriores e tentar prever os próximos pontos da curva (valores das ações). Eles serão descritos a seguir:

- Linear

A primeira função utilizada nesse algoritmo é a linear. Essa é a função mais simples das usadas. Ela tenta encontrar a função

$$f(x) = a \times x + b \quad (1)$$

que melhor aproxima os pontos anteriores, de forma que consiga abranger o máximo de pontos possíveis.

- Polinomial

A segunda função usada já é mais complexa, trata-se da polinomial. Para a função polinomial é possível assumir diversos graus para a função. Como padrão, o algoritmo utiliza a função polinomial de grau dois, mas é possível alterar de maneira fácil para que se obtenha funções com graus diferentes para representar de diferentes formas a curva da função polinomial. Esse tipo de função se constrói passando por um número  $n$  de pontos, onde:

$$n = \text{Grau} + 1 \quad (2)$$

- Base Radial

A terceira função é a Base Radial. A base radial tenta produzir uma função do tipo

$$f(x) = e^{(||x-x'||^2*Y)} \quad (3)$$

onde  $Y$  uma constante de escolha de quem usá-la. Esse tipo de função se mostrou melhor para aproximar funções a partir dos dados, e nesse trabalho será a maior referência para as previsões que serão feitas.

### 3. Análise dos Resultados

Para a analisar os resultados desse algoritmo, foram usados valores dos preços das ações do Facebook no mês de Novembro de 2019. Os valores das ações e os dias em que esses valores ocorreram podem ser observados na Tabela 1.

**Table 1. Valores das ações do Facebook.**

Dia	Valor da Ação	Dia	Valor da Ação
01/11/2019	192,85	14/11/2019	192,92
04/11/2019	194,55	15/11/2019	194,25
05/11/2019	195,36	18/11/2019	194,55
06/11/2019	194,02	19/11/2019	197,39
07/11/2019	191,91	20/11/2019	198,58
08/11/2019	190,00	21/11/2019	197,41
11/11/2019	189,92	22/11/2019	198,38
12/11/2019	190,00	25/11/2019	199,52
13/11/2019	194,69	26/11/2019	200,00

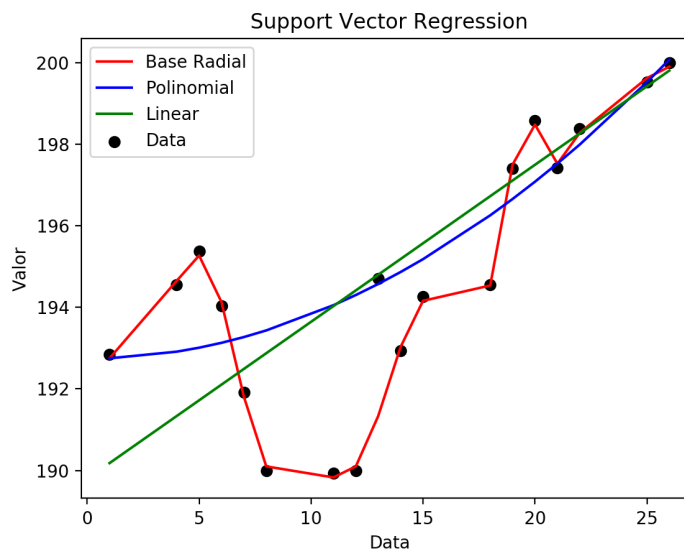
Com estes dados, o algoritmo foi executado e tentou-se prever o valor da ação em alguns dias e compará-los com os resultados reais dos valores. Abaixo, a Tabela 2 mostra o valor das ações previstas pelas funções linear, polinomial, base radial e por último o valor real.

**Table 2. Previsões do algoritmo para as ações do Facebook.**

Dia	V. Linear	V. Poli.	V. B. Rad	Valor Real
27/11/2019	202,58	200,62	200,19	199,90
29/11/2019	203,69	201,85	200,96	201,60

Como é possível observar, a função linear não se mostrou precisa na previsão dos valores futuros. Deve ser utilizada para entender se, de modo geral, o preço de uma ação está caindo ou descendo e em que intensidade isso ocorre.

Já as funções polinomiais e de base radial produziram resultados de forma satisfatória. No primeiro caso o valor da função da base radial ficou um pouco mais próxima do valor real e no segundo caso, o valor da polinomial se aproximou mais. Abaixo, uma imagem que mostra as curvas traçadas pelo algoritmo para a previsão do dia 27.



**Figure 1. Resultado do algoritmo para o dia 27/11/2019**

Uma análise do gráfico permite observar que o modelo Base Radial, em vermelho, representa melhor uma curva que aproxima os pontos anteriores aos previstos, isso explica porque ela foi bem sucedida na previsão dos pontos futuros. Já a função polinomial foi bem sucedida em representar os pontos na faixa superior aos valores iguais a 192. Como as previsões resultaram em valores maiores que estes, também foram bem representadas pela função polinomial.

Já a função linear pouquíssimo representou os pontos anteriores. Ela apenas mostra o padrão de crescimento dos valores com uma taxa de crescimento de aproximadamente \$0,4 por dia. Como ela não representa tão bem os pontos anteriores, seria difícil que ela representasse bem os pontos futuros.

## 4. Bibliografia

### References

- ITNEXT. Disponível em: <https://itnext.io/facebook-stock-prediction-bcfc676bc611/>. Acesso em: 29 Novembro 2019.
- Machine Curve. Disponível em: <https://www.machinecurve.com/index.php/2019/09/20/intuitively-understanding-svm-and-svr/>. Acesso em: 02 Dezembro 2019.
- Awad M., Khanna R. (2015) Support Vector Regression. In: Efficient Learning Machines. Apress, Berkeley, CA. Disponível em: [https://link.springer.com/chapter/10.1007/978-1-4302-5990-9\\_4](https://link.springer.com/chapter/10.1007/978-1-4302-5990-9_4). Acesso em: 02 Dezembro 2019.
- Towards Data Science. Disponível em: <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>. Acesso em: 02 Dezembro 2019.