

# Zoom Preditivo: Predição de Interação na Visualização de Dados na I4.0 Utilizando o Princípio de Inovação

Luis A. Z. Pozenato<sup>1</sup>, Rodrigo da R. Righi<sup>1</sup>, Cesar D. P. Crovato<sup>1</sup>,  
Cristiano A. da Costa<sup>1</sup>, Vinicius F. Rodrigues<sup>1</sup>

<sup>1</sup>PPG em Computação Aplicada - UNISINOS - Av. Unisinos, 950, São Leopoldo, RS  
lpozenato@edu.unisinos.br, {rrrighi, ccrovato, cac, vfrodrigues}@unisinos.br

**Abstract.** *The fourth industrial revolution, called Industry 4.0, has a profound impact on both producers and consumers through massive data generation. Current data analysis methods act reactively to users' interaction, and, consequently, they are unable to provide rapid availability of information. The present work proposes a predictive model to reduce the processing and response time to visualize this large volume of data. Using the concepts of divide and conquer and innovation, together with entropy measures, the proposed model proactively identifies intervals to be explored in temporal data.*

**Resumo.** *Atualmente, a quarta revolução industrial, chamada de Indústria 4.0, traz um profundo impacto tanto para os produtores quanto para os consumidores através da geração massiva de dados. Métodos atuais de análise de dados atuam reativamente à interação dos usuários e, conseqüentemente, não oferecem rápida disponibilidade de informação. O presente trabalho propõe um modelo preditivo para reduzir processamento e tempo de resposta para exploração visual deste grande volume de dados. Utilizando os conceitos de divisão e conquista e de inovação, juntamente com medidas de entropia, o modelo proposto proativamente identifica intervalos para serem explorados em dados temporais.*

## 1. Introdução

A terceira revolução industrial iniciada em 1970 que trouxe grandes avanços para a sociedade, agora cede lugar a um novo modelo de indústria, a indústria 4.0 [Lasi et al. 2014, Lu 2017, Zhou et al. 2016]. A indústria 4.0 tem um relacionamento estreito com os chamados Sistemas Cyber Físicos (CPS) [Zhou et al. 2016, Lu 2017, Xu and Duan 2019], que são sistemas físicos cuja operação é controlada e monitorada por um centro de comunicação, através de sensores, atuadores e processadores [Rajkumar et al. 2010, Ribeiro 2017]. Em conjunto com CPSs, o emprego de Internet das Coisas possibilita uma grande quantidade de dispositivos e sensores gerarem uma massa de dados consideravelmente grande. Como resultado, esse ecossistema provê um conjunto de dados heterogêneos massivo, ao qual geralmente é referenciado por *BigData* [Xu and Duan 2019]. Além de muitas vantagens, através da análise dessa massa de dados é possível identificar pontos fracos nos processos e sistemas com a descoberta de padrões e detecção de anomalias. Porém, tecnologias para lidar com esses dados se tornam desafiadoras.

Atualmente, é notável o auxílio de ferramentas que permitem a análise visual dos dados. Igualmente, trabalhos como o de [Shneiderman and Plaisant 2019], deixam mais evidente que uma das chaves para essa análise com eficiência se encontra em ferramentas visuais que suportem a rápida interação para exploração visual de dados. Porém esse

tipo de ferramentas tende a ter problemas com a demora para resposta na interação do usuário, ou com o grande consumo de processamento [Im et al. 2013]. Esforços estão sendo conduzidos no sentido de diminuir o tempo de transmissão da informação, como o trabalho de [Nedel1 et al. 2019], que introduz um modelo que consegue compactar a informação para transmissão sem grandes perdas, e o trabalho de [Im et al. 2013] que otimiza a distribuição de processamento, trazendo uma resposta visual incremental. Esses modelos representam melhorias significativas, porém, observando-os, conseguimos perceber a atuação do modelo de forma reativa à interação do usuário na ferramenta, o que poderia ser otimizado se houvesse um tratamento preditivo.

Neste contexto, utilizando os conceitos de Divisão e Conquista e de Inovação [Duque et al. 2005], juntamente com o conceito de Entropia [Shannon 1948] e aproveitando a lacuna apresentada anteriormente na atuação de forma preditiva na análise visual de dados, o presente trabalho propõe um modelo que identifica, de forma preditiva, as possíveis interações para exploração na visualização de dados temporais. Empregando o conceito de inovação, o método calcula a entropia de pequenos conjuntos de dados temporais de um dataset para determinar janelas de tempo de interesse. Para tal, a entropia é calculada através do método de Rényi. Em suma, a ideia é mostrar um determinado gráfico e já determinar áreas de interesse nesse gráfico e já trazê-las para a máquina cliente, de modo que se for dado um zoom na área em particular, instantaneamente um novo gráfico com detalhes são apresentados. Tal atitude melhora a experiência na análise eficiente de dados, tarefa que é muito requerida nas áreas da Indústria 4.0, bolsa de valores e Internet das Coisas. Foi desenvolvido um protótipo cujos resultados demonstram a capacidade do modelo de identificar zonas temporais de interesse.

O restante deste trabalho se encontra organizado em 5 seções. A Seção 2 apresenta os trabalhos relacionados pesquisados na literatura, enquanto a Seção 3 introduz o modelo proposto. Em seguida, as Seções 4 e 5 descrevem a metodologia de avaliação e os resultados obtidos, respectivamente. Por fim, a Seção 6 conclui este trabalho.

## **2. Trabalhos Relacionados**

Como base para o este trabalho, se faz necessário saber quais são os métodos destacados na literatura para detecção de eventos em dados temporais. Com isso em mente, esta seção descreve algumas abordagens existentes na literatura. [Shneiderman and Plaisant 2019] foca na análise visual de eventos e nos ajudam a reforçar a importância desse tipo de análise, destacando o estado atual da arte. Esse trabalho inclui o estudo de caso de duas ferramentas distintas para análise de eventos temporais. O estudo de casos deixa evidente que o maior benefício das ferramentas atuais é em relação ao aumento da capacidade de análise que elas nos propiciam, conseqüentemente, aumentando a nossa capacidade de lidar com a complexidade do mundo real, com tarefas como limpeza de dados, descoberta de padrões que estejam sendo usados da forma errada, ajuste de contexto dos dados antes da sua utilização e identificação de registros de interesse. No trabalho desenvolvido por [Im et al. 2013], os autores apontam dois problemas em ferramentas de análise visual: a demora na resposta da interação do usuário e/ou o grande consumo de processamento, apresentando uma abordagem para visualização interativa de grandes conjuntos de dados, tendo, como principal característica o retorno incremental ao usuário. Porém não deixam claro como a interação para exploração de forma ágil foi conduzida.

O trabalho desenvolvido por [Gabarda and Cristóbal 2010] nos traz uma boa visão sobre a utilização de entropia para a detecção de eventos em séries temporais. Os autores utilizam a entropia de Rényi como ferramenta para localizar eventos de atividade sísmica aplicado a leituras de sinais de terremotos e atividades vulcânicas. Usando dados reais e sintéticos, o método facilita a detecção de eventos utilizando a amplitude e a frequência de um sinal e tem o objetivo de prover informações valiosas sobre eventos em sinais de tempo discreto, através da detecção de eventos de interesse. Em [Duque et al. 2005], os autores utilizam o conceito de divisão e conquista juntamente com o princípio de inovação para criar um modelo que detecta distúrbios na qualidade de energia elétrica. O princípio de inovação é encontrado na formulação do filtro de Kalman e indica a quantidade de informação não presente em uma amostra anterior. Por fim, [Nedell et al. 2019] nos trazem uma boa visão dos motivos da compactação da informação de dados temporais ao ser transmitida por um meio de comunicação como a Internet. O volume de dados *BigData*, se for transmitido na íntegra, demandaria bastante tempo, além de não poder ser representado na sua totalidade em dispositivos de saída. O modelo adotado por [Nedell et al. 2019] divide a amostra de dados em partes menores, identifica os valores mínimo e máximo de cada amostra e transfere somente estes valores pela rede de comunicação.

A Tabela 1 apresenta um comparativo dos cinco trabalhos descritos acima mostrando suas lacunas e oportunidades observadas. A tabela referencia cada trabalho como segue: (TR1) [Shneiderman and Plaisant 2019]; (TR2) [Im et al. 2013]; (TR3) [Gabarda and Cristóbal 2010]; (TR4) [Nedell et al. 2019]; e (TR5) [Duque et al. 2005]. A visualização de dados temporais que suporte a rápida interação para exploração, é um benefício que trará a agilidade necessária na indústria 4.0. As ferramentas, no entanto, tendem a ter problemas quanto à demora para resposta na interação do usuário e/ou um grande consumo de processamento. Analisando os trabalhos atuais na literatura, verifica-se que a forma de atuação adotada para sanar esses problemas é de forma reativa à interação do usuário. Utilizar processos de detecção de eventos em dados temporais, com a finalidade de detectar inovação de forma preditiva, abre possibilidades para que os sistemas explorem inatividade da interação do usuário para melhor agilidade na resposta.

### **3. Zoom Preditivo**

Tendo em mente os problemas destacados e os benefícios que a análise visual de dados temporais representam, a presente seção descreve o modelo Zoom Preditivo.

#### **3.1. Decisões de projeto**

O modelo pretende destacar visualmente as áreas que podem ser de interesse para a exploração em um conjunto de dados temporais. Conforme destacado por [Gabarda and Cristóbal 2010] quando, analisando eventos em dados temporais, o evento de interesse depende da interpretação do contexto que é utilizado. Para o desenvolvimento do projeto, foram adotadas algumas decisões que guiaram este estudo. Em resumo, as decisões são quatro:

- O modelo busca destacar visualmente as áreas que podem ser de interesse (“Janela de Interesse”) para a exploração em um conjunto de dados temporais;
- A entropia de Rényi é utilizada no modelo para representar quantitativamente a importância da informação presente em uma janela;

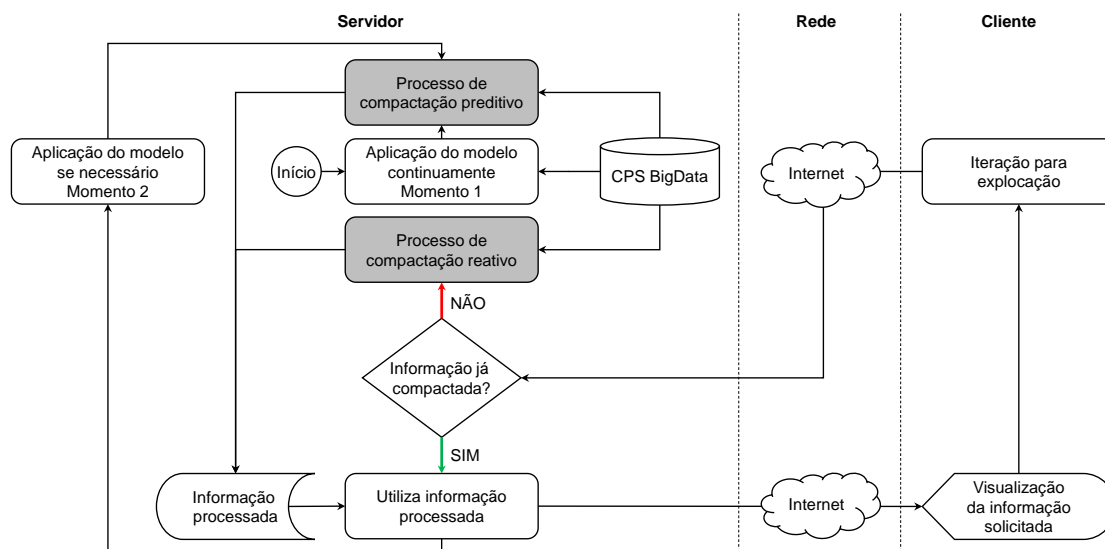
**Tabela 1. Comparativo dos trabalhos relacionados**

Trabalho	Lacunas	Oportunidades
TR1	Estudo da literatura apontando os benefícios de ferramentas que permitem a interação para exploração e estudo de casos de ferramentas para visualização de dados destacando deficiências e benefícios.	Reforça a necessidade de ferramentas que integram a ágil interação para exploração.
TR2	Solução que permite executar pesquisas em dados genéricos, escalável, com resposta incremental que permite a visualização da pesquisa em andamento.	O processo é executado de forma reativa à interação do usuário. A implementação não é descrita. Não fica claro o que é feito para deixar os nós de processamento "aquecidos". Destaca os problemas com a demora na resposta da interação do usuário e/ou um grande consumo de processamento.
TR3	Utilização de entropia para localizar eventos de atividade sísmica em dados temporais.	Demonstra a eficiência da utilização da entropia de Rényi, na detecção de eventos destacada pela literatura.
TR4	Processo para a compactação de informação de dados temporais para transmissão sem grandes perdas, diminuindo o tempo de espera. Atua de forma reativa à ação do usuário.	Eficiente ao reduzir o tempo envolvido na transmissão de dados. Não destaca nenhum processo para redução de tempo de processamento.
TR5	Deteção de eventos de distúrbios de sinal elétrico em tempo real com eficiência.	A utilização dos princípios de inovação e de divisão e conquista que, além de detectarem os eventos com eficiência, permitem a separação do trabalho em partes menores que podem ser exploradas para diminuir os tempos envolvidos em processamentos.

- O modelo aqui desenvolvido utiliza os conceitos de divisão e conquista e de inovação para detectar as janelas de interesse de acordo com [Duque et al. 2005];
- O processo predição acontece sobre um ambiente cliente-servidor onde um cliente interage com uma ferramenta com o objetivo de explorar visualmente dados temporais.

### 3.2. Arquiteutura

A Figura 1 demonstra a arquitetura do tipo cliente-servidor que o modelo emprega. Dividida em três colunas: cliente, rede e servidor a imagem ilustra o processo de interação do cliente, a latência das redes nos processos de comunicação e o trabalho executado pelo servidor. No contexto da arquitetura de computadores cliente-servidor, um usuário, utilizando uma ferramenta de exploração visual de dados, é considerado o cliente.



**Figura 1. Modelo aplicado na arquitetura Client-Servidor.**

O modelo acrescenta uma nova capacidade ao servidor, utilizando predição. Quando a requisição de clientes chega no servidor, este, em um primeiro momento, verifica se a informação requisitada já está processada e, caso já esteja, simplesmente retorna a informação para o cliente. A capacidade de predição dada à arquitetura cliente-servidor é possível graças à aplicação do modelo. Na figura são destacadas duas rotas distintas no servidor: a rota em vermelho e a rota verde. A rota em vermelho representa o processamento no modelo tradicional reativo, em que, para cada requisição de informação, a única saída é processar toda a informação sob demanda para então retornar ao cliente. A rota em verde traz a agilidade necessária na resposta à interação do cliente. Ao receber a requisição de informação o servidor já tem o trabalho de processamento pronto e simplesmente retorna a informação ao cliente. Ter a informação já processada sob demanda, possui um alto custo computacional se aplicado a todo o volume de dados *BigData*, portanto prever qual informação será requisitada permite o processamento apenas do que é necessário.

A aplicação do modelo pode acontecer em dois momentos. Quando os dados *BigData* são gerados, o modelo já pode ser aplicado, dividindo os dados em partes menores no momento em que são gerados e já identificando as janelas de interesse. Em um segundo momento, quando um cliente solicita um intervalo ao servidor, estando ou não a informação compactada, o modelo é aplicado neste intervalo em paralelo sem bloquear a resposta ao cliente. Isso acontece porque o intervalo solicitado contém sub-intervalos que podem ser explorados. Enquanto o cliente estiver analisando a informação visual, o servidor estará se aproveitando da ociosidade, atuando de forma preditiva para que, quando a próxima requisição chegar, o processamento já esteja concluído. Se o sub-intervalo já tiver sofrido a aplicação do modelo então, o segundo momento é ignorado. Identificando as janelas de interesse, o servidor pode atuar em processamentos das janelas identificadas com processamento de compactação. A arquitetura cliente servidor aqui adotada permite uma grande redução de processamento, principalmente em casos que contenham mais de um cliente, visto que a informação processada por um fica disponível para outros.

A Figura 2 apresenta a arquitetura do modelo formada por duas etapas. A primeira parte da arquitetura é a etapa de divisão onde é empregado o conceito de divisão e conquista com o objetivo de diminuir o trabalho em partes menores, facilitando o trabalho e permitindo paralelismos computacionais. Tendo como entrada uma base de dados temporal *BigData*, o princípio de divisão e conquista se encarrega de dividir o trabalho em partes menores, janelas, onde, para cada janela, será então aplicado o processo de detecção de inovação. A divisão do trabalho pode acontecer em uma base de dados estática já existente, ou pode ocorrer em tempo real, conforme os dados são criados.

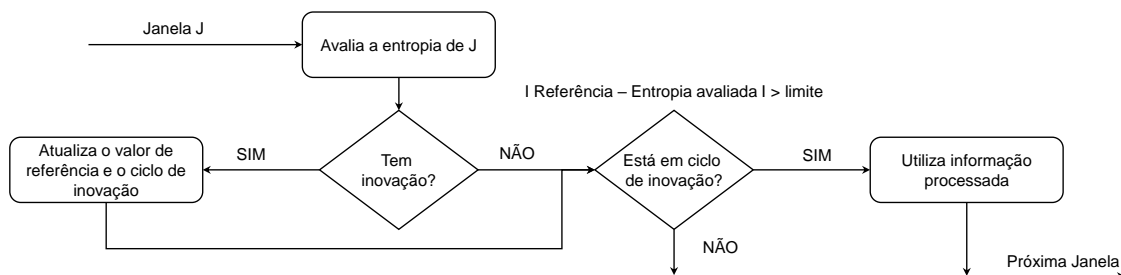


Figura 2. Zoom Preditivo em funcionamento.

A segunda etapa é a parte principal do modelo e é neste momento que é aplicado o conceito de inovação juntamente com o conceito de entropia. Com o objetivo de prever e identificar as possíveis janelas de interesse, esta parte atua sobre cada janela que a divisão e conquista disponibiliza. Nela, cada janela tem sua quantidade de inovação medida e avaliada. O princípio de inovação indica a quantidade de informação não presente em uma amostra anterior. A medida de informação da amostra anterior, na arquitetura, é representada pela variável “referência”. Iniciando com o valor zero, a variável “referência” guarda a medida de entropia da última janela avaliada que tenha inovação o suficiente para ser destacada como janela de interesse.

Para encontrar a medida de inovação presente em uma janela, cada janela tem sua medida de entropia subtraída do valor de referência e, então, o valor absoluto desta diferença indica a medida de inovação presente na janela. A medida de inovação da janela é comparada com um valor limite e, caso este valor limite seja ultrapassado, o entendimento é de que há inovação o suficiente nesta janela para ser considerada uma janela de interesse. Ao identificar que uma janela contém inovação, a variável “referência” é atualizada, bem como é inicializado ou finalizado um ciclo de inovação. A nova medida de referência será utilizada para a medida de inovação das próximas janelas e todas as janelas consequentes serão consideradas janelas de interesse até que o ciclo de inovação se encerre. A detecção de inovação acaba por identificar o início e o fim de algum evento e o ciclo de inovação atua como um estado que fará com que todas as janelas, dentro dos limites de um evento, sejam consideradas como janelas de interesse.

### **3.3. Funcionamento**

O modelo é desenvolvido utilizando um o Algoritmo 1, que considera a aplicação sobre uma entrada de dados estática com tamanho pré definido, porém, conforme detalhado na Seção 3.2, o modelo pode ser aplicado em tempo real conforme os dados estão sendo criados. Ele tem como entrada de informação o conjunto de dados temporais “D”, composto de um campo tempo e outro campo contendo o valor naquele tempo. A saída é um conjunto de dados “W” contendo todas as janelas que apresentem inovação o suficiente para serem consideradas janelas de interesse. Para cada janela, é aplicado o conceito de inovação, que indica a quantidade de informação não carregada em uma amostra anterior, sendo assim, é preciso mensurar a quantidade de informação de cada janela. Para mensurar a quantidade de informação da janela com índice “k” do conjunto “D”, primeiro é medida a entropia da janela “k”, com o auxílio da função “E”. Esta função, definida no Algoritmo 2, recebe como entrada um conjunto “DK” e computa a medida de entropia desse conjunto. A entropia é uma função da distribuição de probabilidades em que a soma das probabilidades é 1. Através da distribuição de probabilidades dos dados da janela, é calculada a entropia de Rényi que, por sua vez, é atribuída à variável “entropy”.

Para que uma dada janela seja considerada como janela de interesse, ela precisa conter suficiente inovação, o que é testado no primeiro desvio condicional “se”. Nesse desvio condicional, o valor absoluto da diferença entre a medida de entropia da janela e o valor de entropia de referência é comparado com o valor de limite de inovação. Quando há inovação suficiente em uma janela “Dk” o valor de referência é atualizado com o medida de entropia da janela “Dk”. Esse novo valor de referência atualizado será utilizado para a verificação de presença de inovação nas janelas seguintes. Nesse momento, também é atualizado a variável “incycle”, iniciando ou terminando um ciclo de inovação, se o ciclo

---

**Algoritmo 1:** Identifica as janelas de interesse.

---

**Entrada:** D (Conjunto de dados temporais), length (Quantidade de subdivisões, janelas), threshold (Valor Limite para detecção de inovação)

**Saída:** W (Conjunto de janelas de interesse identificadas)

```
W ← ∅
reference ← 0
incycle ← false
para k ← 0 até length faça
    entropy ← E(Dk)
    entropy ← (1/length) * entropy
    se |entropy - reference| > threshold então
        | reference ← entropy
        | incycle ← ¬true
    fim
    se incycle = true então
        | W ← W ∪ Dk
    fim
fim
```

---

---

**Algoritmo 2:** Mede a entropia de uma janela empregando Rényi.

---

**Entrada:** D<sub>k</sub> (Conjunto de dados temporais)

**Saída:** entropy (Medida de entropia)

probabilities ← p(D<sub>k</sub>)

entropy ← renyi(probabilities)

---

de inovação estava ativo, inativa, e, se estava inativo, ativa. Portanto, a variável “incycle”, assume os valores de uma variável *booleana*, “verdade” ou “falso”. Enquanto o algoritmo estiver com o ciclo de inovação ativo, todas as janelas serão identificadas e armazenadas como janelas de interesse.

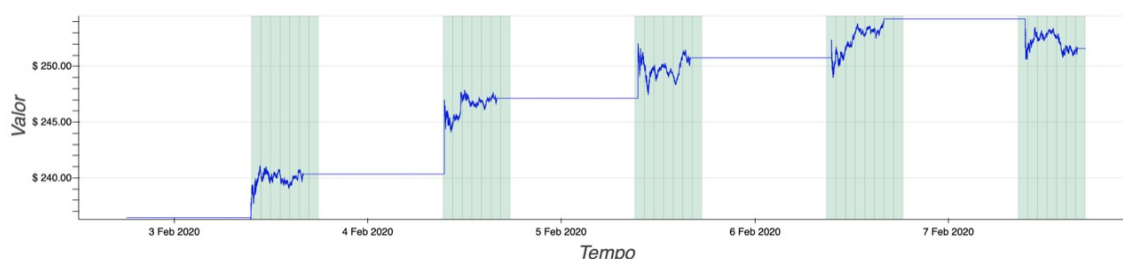
#### 4. Metodologia de Avaliação

Para avaliar o funcionamento do modelo, foi criado um protótipo na linguagem de programação Python. A base de dados utilizada na validação do modelo foi extraída do site kaggle.com e foi encontrada através de uma pesquisa diretamente no Google Datasets. Esta base de dados é pública, está atualmente na revisão 48, se encontra no formato *Comma Separated Values* (CSV), é continuamente trabalhada e há 121 trabalhos fazendo referência a ela na literatura. Tendo o nome de *Intraday Market Data*, os dados contidos na base são referentes ao valor de mercado de ações com intervalo de 3 segundos. O protótipo foi aplicado nesta base de dados sobre a medida de valores de mercado de duas ações distintas, das empresas Apple (www.apple.com) e NVidia (www.nvidia.com), em intervalos semanais, cada intervalo contendo em média 140.000 registros. Como parâmetro para o conceito de divisão e conquista, o protótipo utiliza 100 janelas e, para o limite de inovação utilizado na identificação de janelas, o valor 0,07, conforme apontado por [Duque et al. 2005], para um ambiente com baixo ruído.

#### 5. Análise dos Resultados

Na primeira aplicação do modelo é utilizado o valor de mercado das ações da empresa NVidia, no intervalo de sete dias, das zero horas do dia 1 de fevereiro de 2020 às 24 horas

do dia 7 de fevereiro de 2020. Com um total de 66.762 registros temporais em intervalos de 3 segundos, o resultado da aplicação do modelo é demonstrado na Figura 3. O eixo “x” representa as medidas de tempo, o momento em que foi medido o valor da ação, enquanto o eixo “y”, o valor da ação. As janelas de interesse são destacadas com um retângulo na cor verde claro e a linha da cor azul, a medida do valor das ações no momento. Uma característica logo observada em relação a base de dados de valores de ações é o intervalo entre os dias de trabalho. O mercado de ações possui recessos de um dia para o outro e as oscilações de valores acontecem em horário útil. Este intervalo (linha reta horizontal) de um dia para o outro, ajuda na identificação visual dos períodos de movimentação.



**Figura 3. Valor de mercado das ações da empresa NVidia - fevereiro de 2020.**

A Tabela 2 demonstra as informações medidas nas sete primeiras janelas de interesse que foram destacadas na cor verde na Figura 3. Cada linha desta tabela representa uma janela de interesse. Conforme descrito na metodologia, o limite de medida de inovação para considerar que uma janela é de interesse foi fixado em 0,07. A medida da coluna “Inovação” é o valor absoluto da diferença entre “Entropia” e “Referência”.

**Tabela 2. Janelas de interesse no dia 03-02-2020, NVidia.**

Janela	Início	Fim	Referência	Entropia	Inovação
1	09:34:15	10:45:44	0,000	0,092	0,092
2	10:45:50	11:57:05	0,092	0,085	0,007
3	11:57:11	13:08:23	0,092	0,082	0,010
4	13:08:29	14:19:51	0,092	0,082	0,010
5	14:19:57	15:30:59	0,092	0,081	0,011
6	15:31:05	16:42:31	0,092	0,075	0,016
7	16:42:37	17:53:45	0,092	0,000	0,092

Observando a coluna “Inovação”, verifica-se que a medida de inovação da primeira linha, medindo 0,092, ultrapassa o limite pré estabelecido, portanto esta janela que compreende o intervalo temporal de 09:34:15 a 10:45:44, é considerada como uma janela de interesse. Podemos observar que o valor de referência é zero, o que indica que até então nenhuma janela havia sido encontrada e o valor de entropia da primeira janela passa, então, a ser o valor de referência para as próximas avaliações.

Esta primeira janela de interesse conforme, descrito na Seção 3.2, indica o início de um ciclo de inovação, o que [Duque et al. 2005] descreve como o início de um evento. O modelo ao detectar a primeira janela de interesse, entrará em um estado de ciclo de inovação, tratando as próximas janelas como janelas de interesse, até encontrar uma próxima janela que indique o fim deste ciclo. Assim sendo, as janelas número 2, 3, 4, 5 e 6, mesmo com pouca inovação são identificadas como janelas de interesse. A janela



de número 7, por sua vez, demonstra inovação em relação ao valor de referência que, até o momento, é o da medida de inovação da primeira janela e será substituído pela medida de inovação dessa janela. Isto porque realmente neste intervalo a bolsa de valores estaria encerrando suas atividades, o que é visível pela entropia medida com valor zero e também visualmente no gráfico.

O modelo foi aplicado novamente na base de dados, utilizando os valores de mercado das ações da empresa *Apple*. Os resultados estão disponíveis na Figura 4 e na Tabela 3. Conforme os resultados observados, concluímos que o modelo consegue destacar corretamente as janelas de interesse, identificando assim os intervalos em dados temporais que possivelmente serão explorados. A entropia de Rényi, utilizada para mensurar a medida de informação presente em uma janela, se demonstrou adequada e eficiente. Quanto menor a quantidade de registros em uma janela, menor será a quantidade de informação nela presente e a medida de entropia tende a diminuir. Nesse caso, o limite máximo para detecção de inovação deve ser ajustado. O valor de limite utilizado para detecção de inovação também se demonstrou adequado nos testes realizados.

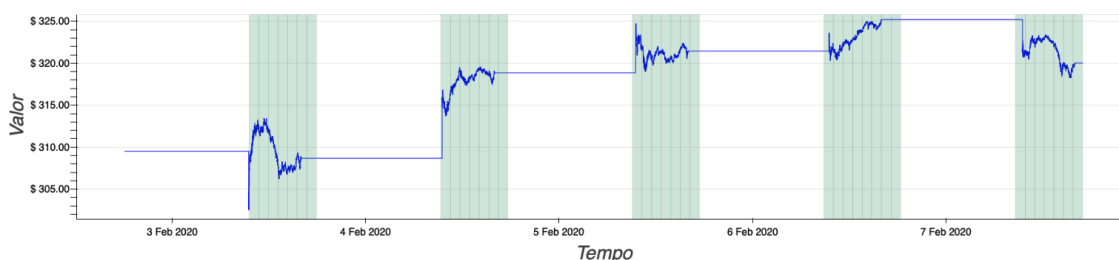


Figura 4. Valor de mercado das ações da empresa *Apple* - fevereiro de 2020.

Tabela 3. Janelas de interesse no dia 05-02-2020, *Apple*.

Janela	Início	Fim	Referência	Entropia	Inovação
15	09:06:32	10:17:43	0,000	0,089	0,089
16	10:17:49	11:29:07	0,089	0,093	0,004
17	11:29:13	12:40:34	0,089	0,091	0,001
18	12:40:40	13:51:49	0,089	0,088	0,001
19	13:51:55	15:03:08	0,089	0,087	0,002
20	15:03:14	16:14:27	0,089	0,087	0,002
21	16:14:33	17:25:28	0,089	0,000	0,089

## 6. Conclusão

A indústria 4.0 já é uma realidade, assim como os desafios da análise do grande volume de dados temporais gerados continuamente por este novo padrão industrial. O modelo aqui apresentado, ao identificar as áreas de interesse para exploração em dados temporais, traz uma abordagem preditiva para confrontar os atuais métodos reativos utilizados em ferramentas que permitem a interação para exploração visual de dados. Zoom Preditivo se demonstra eficiente ao identificar nos dados temporais intervalos que se destaquem pela medida de informação presente na amostra. O modelo permite a sistemas e ferramentas explorarem ociosidades e antecipar processamentos, a fim de mitigar ou diminuir o tempo de espera na interação para a exploração.

Embora eficiente, o Princípio de Inovação baseado na entropia é apenas o ponto de partida desta Pesquisa, sendo necessário analisar outros critérios para determinação de trechos de sinal com grandes chances de zoom. Como trabalhos futuros, critérios como fator de crista, spikes e variações nos valores eficazes podem ser analisados e incluídos no modelo. Inclusive é possível introduzir Machine Learning para identificar esses trechos com zoom potencial. Ainda, fazer ranking dos mesmos é outro passo importante na Pesquisa.

## Agradecimentos

Os autores gostariam de agradecer o suporte das seguintes agências de pesquisa: CNPq, CAPES e FAPERGS.

## Referências

- Duque, C. A., Ribeiro, M. V., Ramos, F. R., and Szczupak, J. (2005). Power quality event detection based on the divide and conquer principle and innovation concept. *IEEE Transactions on Power Delivery*, 20(4):2361–2369. Cited By :28.
- Gabarda, S. and Cristóbal, G. (2010). Detection of events in seismic time series by time-frequency methods. *IET Signal Processing*, 4(4):413–420. Cited By :26.
- Im, J. ., Villegas, F. G., and McGuffin, M. J. (2013). Visreduce: Fast and responsive incremental information visualization of large datasets. In *Proceedings - 2013 IEEE International Conference on Big Data, Big Data 2013*, pages 25–32. Cited By :21.
- Lasi, H., Fettke, P., Kemper, H. ., Feld, T., and Hoffmann, M. (2014). Industry 4.0. *Business and Information Systems Engineering*, 6(4):239–242. Cited By :892.
- Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. *Journal of Industrial Information Integration*, 6:1–10. Cited By :546.
- Nedell, E., Begnini, M., Melchiades, Vinicius, L., Schreiber<sup>1</sup>, Paredes, C. D., Crovato<sup>1</sup>, and Righi<sup>1</sup>, R. d. R. (2019). Compressiot: an optimized compression model for displaying a high volume of iot data in web environments. *Computer*.
- Rajkumar, R., Lee, I., Sha, L., and Stankovic, J. (2010). Cyber-physical systems: The next computing revolution. In *Proceedings - Design Automation Conference*, pages 731–736. Cited By :982.
- Ribeiro, L. (2017). Cyber-physical production systems’ design challenges. In *IEEE International Symposium on Industrial Electronics*, pages 1189–1194. Cited By :16.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(4):623–656. Cited By :19081.
- Shneiderman, B. and Plaisant, C. (2019). Interactive visual event analytics: Opportunities and challenges. *Computer*, 52(1):27–35. Cited By :1.
- Xu, L. D. and Duan, L. (2019). Big data for cyber physical systems in industry 4.0: a survey. *Enterprise Information Systems*, 13(2):148–169. Cited By :75.
- Zhou, K., Liu, T., and Zhou, L. (2016). Industry 4.0: Towards future industrial opportunities and challenges. In *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2015*, pages 2147–2152. Cited By :265.