**Documentação Complementar**

De forma geral, o seguinte fluxo foi utilizado para a implementação

Interface gráfica do usuário, Texto, Aplicativo

Descrição gerada automaticamente

Figura 1: Fluxo de implementação utilizado.

**Análise de Requisitos**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RF1: *Upload* de arquivos | | **Oculto ()** | | |
| **Descrição:** O sistema deve permitir o usuário realizar o *upload* de arquivos | | | | |
| **Requisitos não funcionais** | | | | |
| **Nome** | **Restrição** | **Categoria** | **Desejável** | **Permanente** |
| RNF 01 | O formato de arquivo para *upload* deverá ser *csv* |  | **(x)** | **()** |
| RNF 02 | O sistema deve permitir a visualização do nome de cada um dos atributos do arquivo realizado o *upload.* |  | **(x)** | **()** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RF2: Visualização dos dados de *upload* | | **Oculto ()** | | |
| **Descrição:** O sistema deve permitir a visualização dos dados no formato de *dashboard.* | | | | |
| **Requisitos não funcionais** | | | | |
| **Nome** | **Restrição** | **Categoria** | **Desejável** | **Permanente** |
| RNF 01 | Deve ser exibido para o usuário um conjunto de 3-4 painéis com informações visuais. |  | **(x)** | **(x)** |
| RNF 02 | As visualizações implementadas devem levar em conta aspectos de acessibilidade, como os relacionados ao daltonismo, no uso das cores, por exemplo. |  | **(x)** | **(x)** |
| RNF 03 | As visualizações implementadas devem refletir para o usuário informações sobre a presença de falhas e/ou anomalias identificadas nos dados. |  | **()** | **()** |
| RNF 04 | As visualizações implementadas devem possuir recursos de interatividade para o usuário. |  | **(x)** | **(x)** |

**Design do Dashboard**

Após o levantamento inicial dos requisitos do sistema, o próximo passo consistiu na proposta de uma interface para a ferramenta. No primeiro momento, realizou-se uma consulta por *dashboards* e/ou painéis já existentes, buscando assim por tendências e oportunidades de inovação. No fim, o seguinte *wireframe* foi obtido:

**Gráfico, Gráfico de mapa de árvore

Descrição gerada automaticamente**

Figura 2: Wireframe da interface proposta

**Implementação**

A ferramenta proposta é um sistema *desktop*, desenvolvido em *Python.* Para sua implementação, foram utilizadas as seguintes bibliotecas:

* PyQt5: implementação da interface *desktop*
* Plotly: Implementação das visualizações interativas

A versão final implementada pode ser vista a seguir:

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Figura 3: Versão final do sistema implementada.

**Testes**

**Análise das Visualizações Propostas**

**Características Gerais**

De forma geral, a biblioteca adotada para representação dos gráficos permite a interação analítica da seguinte maneira:

* Destacar (*Brushing and Linking*) = Ao clicar sobre os pontos representados em tela, o usuário consegue verificar as informações referentes aos eixos x e y. Além disso, é possível desenhar um *box select*, ou seja, um retângulo em tela, focando nos dados contidos nesse retângulo e desfocando os demais, conforme a Figura seguinte.

Interface gráfica do usuário

Descrição gerada automaticamente

Figura 4: Exemplo de aplicação da funcionalidade de box select.

* Zooming and panning = Pontos específicos dos gráficos podem ser visualizados de forma mais próxima (*zoom in*) ou distante (*zoom out*). Além disso, é possível acessar áreas que não são visíveis por padrão, através do *panning.* Vale ressaltar que, quando é aplicado o *zoom*, as escalas são alteradas e adequadas conforme a nova dimensão. Um exemplo pode ser visto na Figura seguinte.

Interface gráfica do usuário, Gráfico, Gráfico de dispersão

Descrição gerada automaticamente

Figura 5: Aplicação do "zooming and panning".

* Detalhes sob demanda = Algumas informações são ocultas nos gráficos e acessadas somente quando o usuário arrasta o cursor do *mouse* sobre eles. Assim, conforme é necessário, o usuário pode acessar tal informação. A Figura seguinte tenta exemplificar um caso em que isso acontece.
* Gráfico

  Descrição gerada automaticamente

Figura 6: Informações referentes aos quartis do boxplot sendo acessadas sob demanda do usuário.

**Cores**

Durante a implementação do sistema, aspectos de acessibilidade foram levados em conta. Ao longo das aulas, percebeu-se que a escolha das cores em uma visualização não é e não pode ser totalmente aleatória. Ela deve ser guiada por alguns princípios, sendo alguns deles utilizados nesse trabalho, tais quais:

* Uso de cores saturadas ao invés das cores puras.
* Uso de cores próximas que possuem contraste entre si.
* Considerar aspectos sobre o daltonismo (Protanopia, Tritanopia e Deuteranopia)

Sendo assim, segue abaixo o resultado do sistema submetido a um simulador de daltonismo, considerando as suas três variações mais comuns:

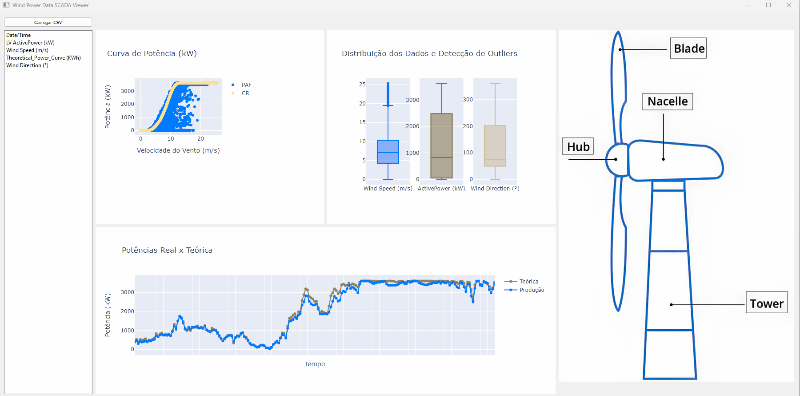


Figura 7: Imagem do sistema submetido ao simulador de daltonismo para o vermelho (Protanopia)

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Figura 8: Imagem do sistema submetido ao simulador de daltonismo para o azul (Tritanopia)

Interface gráfica do usuário, Aplicativo, Word

Descrição gerada automaticamente

Figura 9: Imagem do sistema submetido ao simulador de daltonismo para o verde (Deuteranopia)

**Gráfico de Pontos**

Para a representação da Curva de Potência dos dados coletados, a estratégia padrão utilizada é sempre o uso do scatterplot, considerando o eixo x como sendo a velocidade do vento e o eixo y como sendo a potência produzida. Como o conjunto de dados contém um total de 50.530 valores coletados, é clara a sobreposição de tais pontos em algum momento. Para mitigar esse tipo de ocorrência, adotou-se as seguintes estratégias:

* Redução do tamanho de plotagem
* Remoção da cor de preenchimento
* Uso de cores distintas para

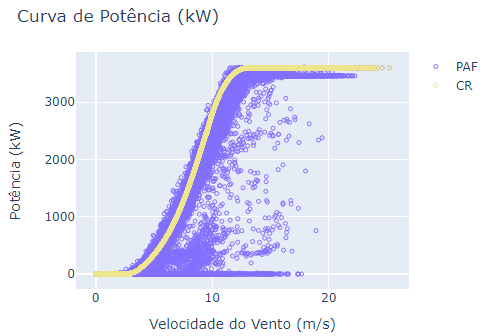


Figura 10: Curva de potência dos dados coletados.

Na Figura 7 acima, a curva em formato de “S”, representada em amarelo, indica a potência teórica ideal que a turbina deveria produzir. Nesse tipo de gráfico, é buscado uma curva com o mesmo formato, mas gerada a partir dos dados coletados. Percebe-se que tal curva é sim plotada em tela, porém, percebe-se também a existência de um conjunto de pontos que não se encaixam em tal curva. De acordo com a literatura, esses pontos podem ser divididos em 4 categorias, conforme a Figura seguinte, e representam dados anômalos, obtidos em momentos de falhas e/ou outras anomalias da turbina eólica.

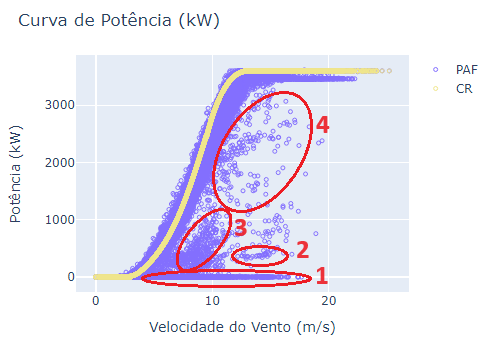


Figura 11: Exemplo de pontos esultants de anomalias e/ou falhas na turbina eólica.

**Boxplot**

O uso desse tipo de visualização permite analisar a dispersão dos valores de interesse, contidos no conjunto de dados, em torno da medida de tendência central definida, nesse caso, a média. Além disso, é possível identificar de forma simples a presença ou não de dados discrepantes, isto é, outliers coletados.

Gráfico, Gráfico de caixa estreita

Descrição gerada automaticamente

Figura 12: Boxplot os principais atributos do conjunto de dados.

**Série Temporal**

Para a representação e comparação das potências real e teórica ao longo do tempo, optou-se por representar tais dados por meio de uma série temporal. Sendo assim, a plotagem de tais dados permitiu identificar alguns aspectos interessantes, tais quais:

* Variabilidade =
* Taxa de mudança =
* Padrões e exceções = pontos + linha

Gráfico, Histograma, Gráfico de dispersão

Descrição gerada automaticamente

Tendência/ pontos + linhas (ver tendência nos dados ao longo do tempo e comparar valores específicos)

Diagrama

Descrição gerada automaticamente

Interface gráfica do usuário

Descrição gerada automaticamente com confiança baixa