

RELATÓRIO: REQUISITOS

1. Estrutura do Projeto (Macrovisão)

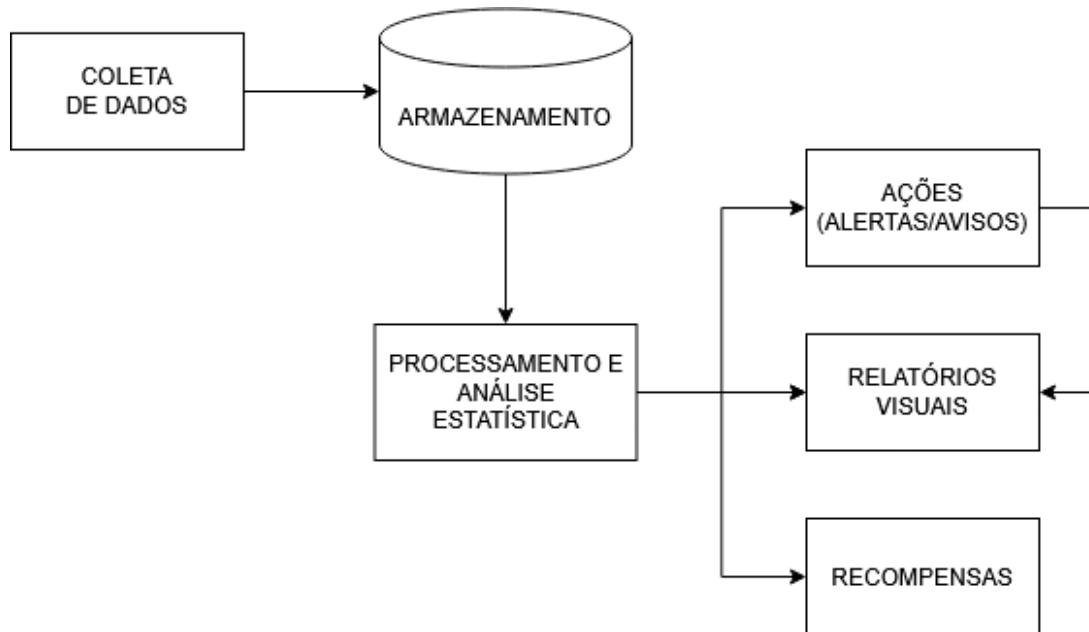


Figura 1 - Diagrama de Integração do Sistema

O projeto proposto, caracterizado como um sistema de monitoramento de sustentabilidade pessoal, tem como objetivo proporcionar uma ferramenta que possibilite a conscientização e o incentivo da sustentabilidade individual por meio do acompanhamento e análise dos seguintes parâmetros: consumo de água, consumo de energia elétrica, geração de resíduos não recicláveis e uso de transporte. Na figura 1 estão explicitos os módulos propostos, sendo estes:

- Coleta de Dados: uma interface para entrada de dados por responsabilidade do usuário.

- Armazenamento: banco de dados relacional que irá guardar todos os dados coletados, dados processados, alertas/avisos, recompensas, métricas, datas e horários de cada usuário.

- Processamento e Análise Estatística: responsável pelo processamento e normalização dos dados brutos, utilizando de técnicas de estatísticas e de machine learning.

- Ações: responsável por correlacionar as informações tratadas e promover ações, como alertas e avisos, de acordo com o nível de sustentabilidade representados.

- Relatórios Visuais: dashboard interativa com gráficos gerais e gráficos contidos em relatórios periódicos.

- Recompensas: responsável por dar propósito e incentivo à sustentabilidade, por meio de gamificação, microtransações, etc.

2. Interação entre os Parâmetros

A integração dos parâmetros será feita em duas etapas: normalização e ponderação final, visando a geração de um índice de sustentabilidade balanceado e ajustado aos fatores externos de cada usuário.

Quanto a normalização, cada parâmetro impacta o Meio Ambiente de forma diferente e com pesos diferentes – um litro de água desperdiçada não é equivalente, em termos de impacto, a um quilo de CO₂ emitido. Paralelamente, fatores externos, como a quantidade de moradores de uma residência, podem levar a dados que não representam o índice individual corretamente. Desta maneira, a etapa de normalização irá considerar os fatores externos para obter métricas individuais sobre os dados e, então, converter cada parâmetro em uma escala comum utilizando do método de normalização **min-max**, representado pela fórmula:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Onde x' é o valor normalizado, x é o valor original, $\min(x)$ é o valor mínimo que x pode assumir e $\max(x)$ é o valor máximo. Utilizando desta normalização, obteremos um valor entre 0 e 1.

Quanto a ponderação final, com os valores normalizados, podemos utilizar a média ponderada e atribuir pesos para cada parâmetro de acordo com o nível de impacto, por exemplo, podemos atribuir um peso maior para a emissão de CO₂, gerado por combustíveis fósseis, e um peso menor para o consumo exacerbado de água. Desta forma, podemos obter o índice de sustentabilidade individual que será a métrica base de todo sistema.

3. Controle dos Parâmetros

Por meio de todas as etapas do sistema, será empregado, por da gamificação, um sistema de recompensas que incentiva os usuários a melhorarem seus hábitos de sustentabilidade; o sistema, também, irá atribuir a cada usuário uma “medalha” que representa seu índice de sustentabilidade individual, esta pode melhorar ou piorar dependendo deste índice.

Por meio das ações e relatórios com gráficos, o sistema irá fornecer um feedback constante para o usuário, com dicas e observações para melhorar a sustentabilidade, efetivamente guiando para uma melhora dos parâmetros ao longo do tempo.

4. Coleta de Dados

O principal meio de coleta dos parâmetros será por entrada manual, realizada pelo próprio usuário, para os parâmetros que suportam esta abordagem (consumo de água e consumo de energia elétrica), será utilizado um sistema de OCR (reconhecimento ótico de caracteres) para promover mais praticidade para o usuário. Entretanto, para parâmetros que não pode ser reconhecidos por meio de simples fotos, será necessário a entrada dos dados em um formulário integrado ao sistema. No caso da coleta do uso de transporte, para cada tipo de meio de transporte será atribuída uma média de emissão de CO₂, para determinar o impacto.

Para cada parâmetro, o método de coleta se dá:

- Consumo de água: inserção manual ou “foto” do hidrômetro;
- Consumo de energia elétrica: inserção manual ou foto do medidor de energia elétrica (“relógio de luz”);
- Geração de resíduo não reciclável: inserção manual do peso dos resíduos não recicláveis e dos resíduos recicláveis (para comparação);
- Uso de transporte: inserção manual do tipo de transporte e distância percorrida.

5. Correlação entre Parâmetros

Para desencadear algumas das ações, é necessário correlacionar os parâmetros com fatores externos e com eles mesmos, de forma a gerar métricas úteis para prever o comportamento do usuário dada certas circunstâncias. Por exemplo, podemos relacionar a época do ano com o aumento da energia elétrica e gerar uma ação para o usuário sobre isto, como um aviso; pode-se, também, relacionar o consumo de energia da semana anterior com o da semana atual, para detectar se houve melhora ou piora de certos parâmetros.

De forma geral, a geração das ações irá se basear de alguma forma na correção dos parâmetros com outros fatores, de forma que um alto índice de sustentabilidade, em geral, promova ações positivas. No entanto, parâmetros que contribuam mais para o índice podem obscurecer parâmetros de menor significancia (na média ponderada), a correlação de anterioridade (o parâmetro com ele mesmo porém em uma data passada) irá permitir detectar tais casos.

6. Estatísticas

Para realização das correlações, será necessário implementar técnicas estatísticas e de *Machine Learning* para interpretação e geração de métricas

úteis, dada a baixa complexidade necessária para o projeto os métodos adotados serão, inicialmente:

- Coeficiente de correlação de Spearman: também chamada de “rô de Spearman”, esta técnica mede o grau de correlação entre duas variáveis não-lineares de escala métrica (contável ou mensurável) e gera um coeficiente entre dois valores -1 e 1, onde quanto mais perto dos extremos, mais correlacionados estão as variáveis.

- Árvore de Decisão: um algoritmo recursivo de Machine Learning utilizado para regressão ou classificação, utiliza de “galhos” com parâmetros para determinar o resultado final. Se utilizada como algoritmo de regressão, gera um valor a partir dos dados fornecidos, já se utilizada como algoritmo de classificação, aponta para uma das categorias pré-determinadas (significa que o dado se encaixa mais em uma categoria / classificação).

Utilizando do coeficiente de Spearman para determinar a correlação entre o valor do parâmetro atual com o valor do parâmetro passado, pode-se observar uma melhora ou piora da estabilidade ao longo do tempo. E utilizando da árvore de decisão para regressão, podemos obter um valor estimado futuro do parâmetro, sendo o tempo determinado pela frequência da coleta de dados.

9. Relatórios

Os relatórios irão conter os gráficos dos períodos que eles representam, ou seja, se um relatório for semanal, o gráfico irá conter todos os dados da semana registrada; serão contidos nos relatórios, também, os alertas e recomendações registrados. O usuário irá poder consultar os gráficos gerais (abrangem todo o período de utilização), também, pela aba de relatórios.

Para cada parâmetro, os gráficos serão:

- Água: Gráfico de Linhas (consumo/tempo);
- Energia Elétrica: Gráfico de Colunas (consumo/tempo);
- Resíduos: Gráfico de Setores (% de resíduos);
- Uso de transporte: Gráfico de Setores (% de meio de transporte);
- Geral: Gráfico de Radar (multivariável);

10. Ações

Para cada patamar de sustentabilidade individual (índice dividido por quantidade de patamares, inicialmente 3), serão geradas notificações e alertas personalizados que informam sobre desvios de padrões, estimativas de

consumo e recomendações. Podem ser anexados às recomendações, infográficos visuais com orientações para melhorar a sustentabilidade.

A cada melhora detectada, serão também conferidas recompensas proporcionais ao nível de melhora do usuário, incentivando mais aqueles que se dedicarão para melhorar seus comportamentos.

11. Bibliografia

WIKIPEDIA. Coeficiente de correlação de postos de Spearman. Disponível em: https://pt.wikipedia.org/wiki/Coeficiente_de_correla%C3%A7%C3%A3o_de_postos_de_Spearman.

PSICOMETRIA ONLINE. O que é correlação de Spearman? Disponível em: <https://www.blog.psicometriaonline.com.br/o-que-e-correlacao-de-spearman/>.

TERA BLOG. Árvores de decisão. Disponível em: <https://blog.somostera.com/data-science/arvores-de-decisao>.

CODECADEMY. Normalization. Disponível em: <https://www.codecademy.com/article/normalization>.