

# Utilizando o *Learning Analytics* com o K-Means para Análise de Dificuldades de Aprendizagem na Educação Básica

Arianne S. Torcate, José Carlos F. Barbosa, Cleyton Mário de O. Rodrigues

Universidade de Pernambuco - *Campus* Garanhuns (UPE)  
CEP 55.294 - 902 – Garanhuns – PE – Brasil

{ariannesarmento0, josecarlosfelix8}@gmail.com,  
cleyton.rodrigues@upe.br

## 5 Análise dos Resultados

Com o intuito de visualizar os dados classificados e agrupados pelo K-Means, utilizou-se dois visualizadores distintos, como já foi mencionado, são eles: Scatter Plot e Silhouette Plot. A relação estrutural entre o algoritmo e os visualizados são expostos na Figura 1.

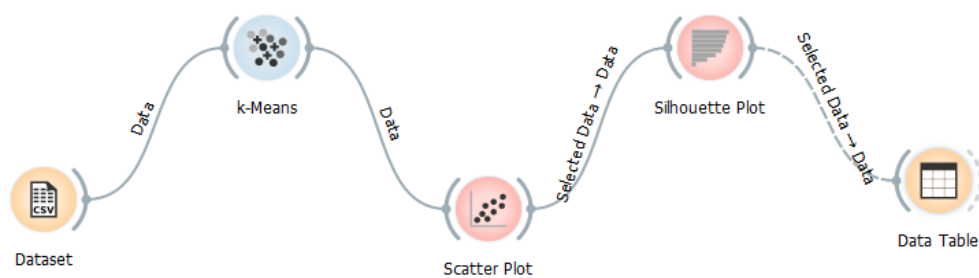


Figura 1. Relação estrutural do K-Means com os visualizadores.

Conforme ilustra a Figura 2, a visualização dos dados através do Scatter Plot nos fornece a seguinte interface:

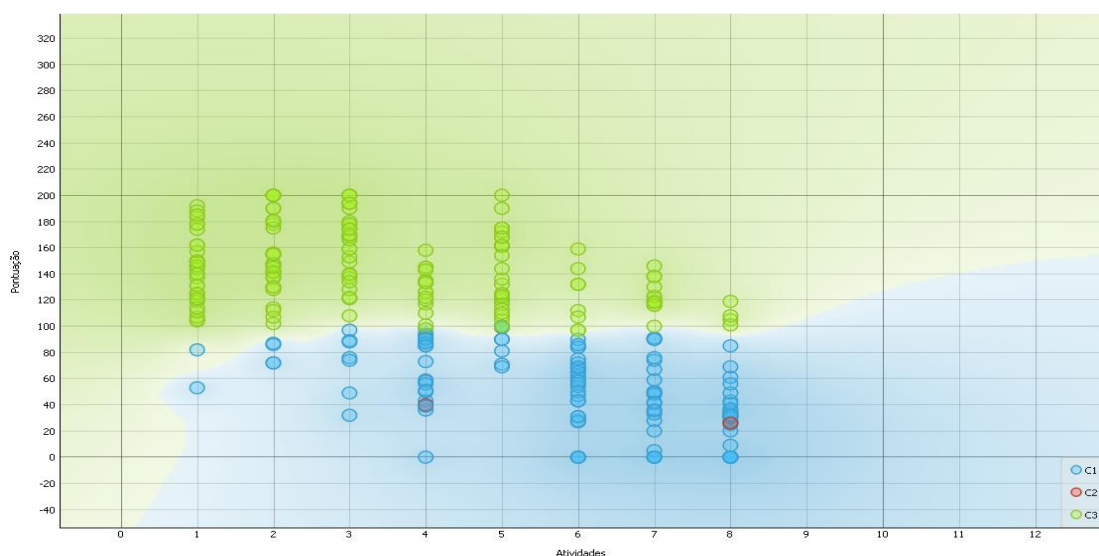
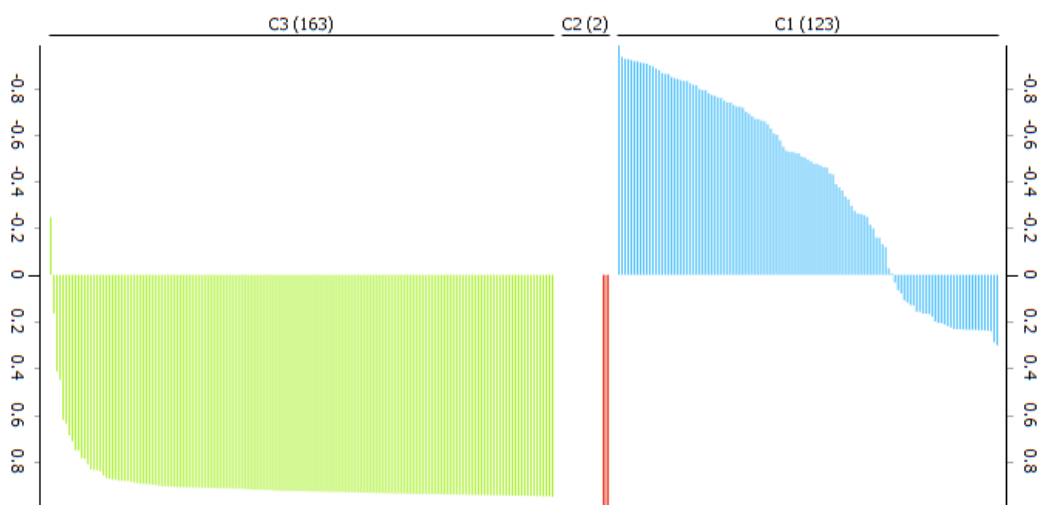


Figura 2. Visualização dos dados através do Scatter Plot.

Na Figura 3 é apresentado a visualização através do Silhouette Plot, que nos fornece uma interface voltada a dados numéricos em relação aos *clusters*.



**Figura 3. Visualização dos dados através do Silhouette Plot.**

Como as Figuras apresentam, através dos visualizadores é possível enxergar os dados de cada *clusters* sob perspectivas diferentes. Lembrando que foram desenvolvidos oito jogos que são correspondentes aos oito conteúdos abordados na pesquisa.

De acordo com o Silhouette Plot, o *cluster* 1 possui dados que em sua grande maioria estão abaixo de 0, isso reforça o que já foi afirmado anteriormente, que esse grupo é composto por alunos que tiveram um desempenho totalmente baixo nos jogos referente aos conteúdos de decimais, operações com frações, expressões numéricas e transformações de unidades de medidas.

Os parâmetros utilizados para chegar a esses resultados foram a pontuação dos alunos por jogo. Mas, no decorrer das análises foi possível identificar que os alunos do *cluster* 1 tiveram um número de tentativas maior e passaram mais tempo para resolver as questões propostas no jogo. Ou seja, o número de tentativas e o tempo são dois parâmetros que colaboram na união e correlação desse *cluster*.

E como já foi afirmado, o *cluster* 2 contém dois alunos que tiveram um número de tentativas exacerbadas para poder conseguir uma boa pontuação nos jogos referente a decimais e transformações de unidades de medidas. Ou seja, esse *cluster* foi criado para agrupar apenas essas duas exceções.

Já no *cluster* 3 é possível identificar que os dados estão acima de 0, ou seja, são dados considerados significativamente bons. Dessa forma, os alunos que tiveram um ótimo desempenho nos jogos referente aos conteúdos de porcentagem, potenciação, as

quatro operações e frações são alunos que também tiveram um número de tentativas baixo e que gastaram menos tempo resolvendo as questões propostas dentro do jogo.

Portanto, é notório que os parâmetros de pontuação por jogo foram âncoras para relacionar esses *clusters*, mas, os parâmetros de tentativas e tempo ajudaram a reforçar ainda mais a correlação entre os dados individuais de cada *cluster*.

Com estas análises fica evidenciado que o objetivo deste trabalho quanto a identificar dificuldades de aprendizagem em oito conteúdos matemáticos no ensino básico foi alcançado com êxito. Além disso, também foi identificado os conteúdos que os alunos apresentaram um ótimo desempenho.

Por fim, utilizar os dois visualizadores de dados em conjunto foi importante para verificar a relação dos dados que estavam sendo analisados. Além disso, enxergar outros parâmetros na análise foi totalmente relevante para confirmar a veracidade das informações de cada *clusters*, visto que ambos foram criados de forma automática pelo algoritmo K-Means no software Orange Canvas.