Resumo 15 – PLN

Na aula de número quinze de Processamento de Linguagem Natural, o professor Jesús mencionou e demonstrou alguns grafos aplicados à obra de Machado de Assis e, para aplicar uma análise mais detalhadas nesses gráficos, é possível aplicar um algoritmo de redução de dimensionalidade, retornando assim características como número de vértices, número de arestas, tamanho, densidade, grau médio, diversidade lexical (número de palavras diferentes do vocabulário que estão presentes na obra), menor caminho médio, entre outros.

Com isso, observa-se um resultado em duas dimensões. Como exemplo, foram coletadas todas as obras de Machado de Assis e plotadas em um gráfico bi dimensional. Para comparação, foram coletadas duas obras de Aluísio de Azevedo e duas obras de José de Alencar para serem plotadas no mesmo gráfico bi dimensional. O resultado obtido foi coerente: é possível perceber uma diferença entre os pontos relacionados a Machado de Assis e os pontos dos dois outros autores, ou seja, os pontos das respectivas obras ficaram visualmente separadas no gráfico. Tais medidas são extremamente importantes, pois servem para a análise dos dados, visualizando as características desses dados, como também na síntese dos dados, identificando alguns padrões característicos daqueles dados e podendo, quem sabe, simular tais características. Imagine que seja possível criar obras com a mesma característica das obras de Machado de Assis.

Um outro ponto extremamente importante abordado durante a aula foi a demonstração de um método criado por uma aluna do professor Jesús. Caminhando para o final da aula, o professor comentou sobre a segunda entrega do projeto final, modificando assim o prazo de entrega e estipulando a data final para domingo.