

Metrificação e análise de carreiras de jogadores de futebol por meio de um sistema computacional

Contexto

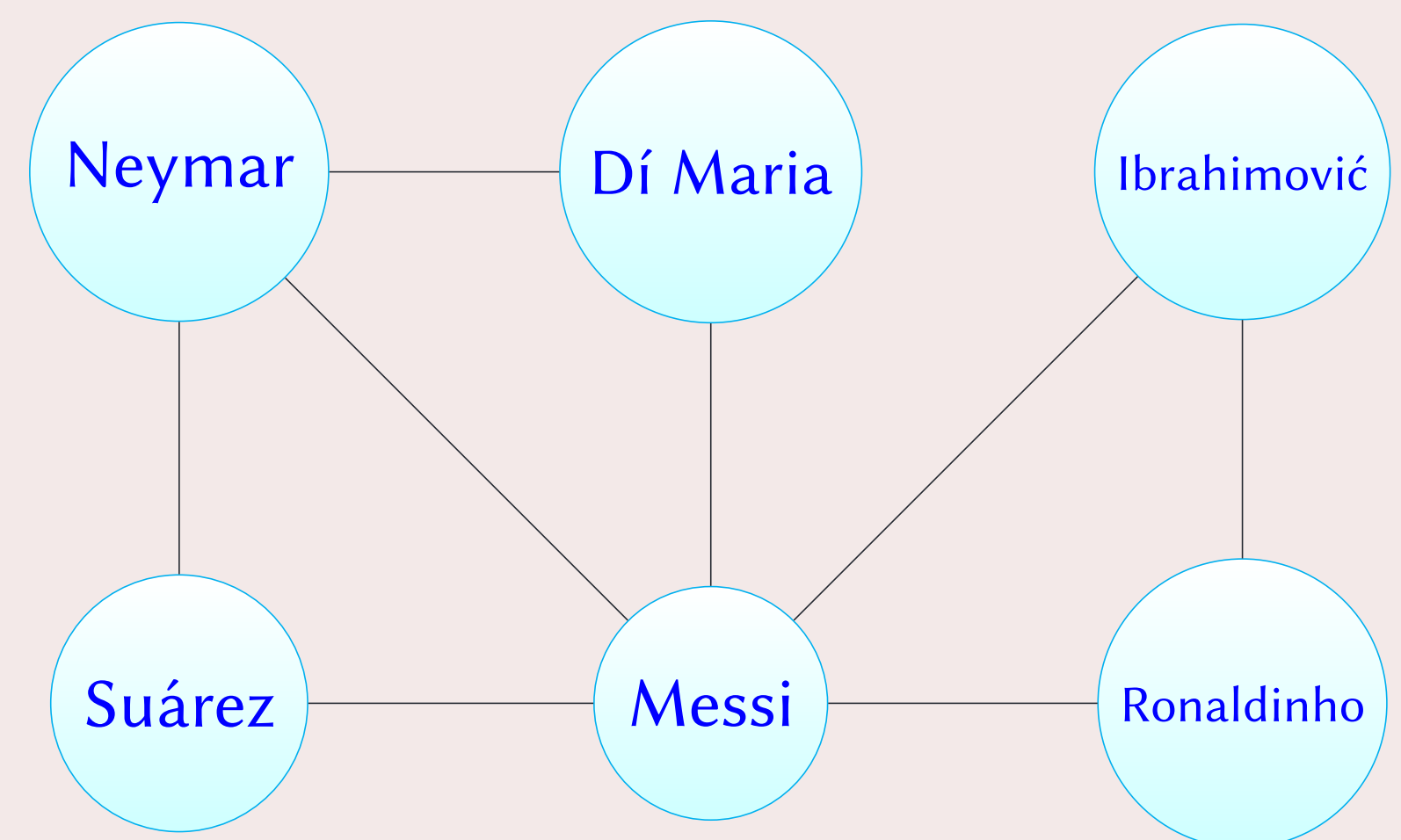
- ▶ Futebol é o esporte mais popular do planeta, e está presente na vida de bilhões de pessoas, dentro e fora do Brasil.
- ▶ Sempre são feitas comparações para decidir os melhores jogadores de cada clube, seleção, temporada e até mesmo da história.
- ▶ Desse contexto, surge a ideia de avaliar a qualidade individual dos atletas. Apesar de já existirem meios para isso, como a Bola de Ouro, seus critérios não são sistemáticos e variam ao longo dos anos.

Resumo

- ▶ Implementação de um sistema que calcula pontuações para as carreiras de jogadores de futebol, baseando-se na fórmula matemática de Salvador et al. (2022).
- ▶ Fluxo:
 - Extrair da Internet dados dos atletas, times e competições;
 - Armazená-las em um banco de dados orientado a documentos;
 - Aplicar a fórmula;
 - Visualizar os resultados.
- ▶ O banco orientado a documentos contém 6 coleções, que representam os registros dos jogadores, times, campeonatos, países, continentes e as pontuações calculadas.
- ▶ Outro componente é um banco de dados orientado a grafos, que permite computar informações envolvendo as relações inter-jogadores, que vão para além das pontuações individualmente calculadas.
- ▶ A ferramenta Apache Superset é utilizada para a visualização dos dados, pois fornece um ambiente de criação de gráficos e tabelas que acompanha a evolução do BD automaticamente.
- ▶ A fórmula foi aplicada em mais de 700 jogadores, contando também com cerca de 1000 times e 500 campeonatos.

Estrutura do Grafo

- ▶ Cada vértice é um jogador e cada aresta representa a atuação conjunta de dois atletas em um mesmo campeonato pelo mesmo clube.
- ▶ Exemplo de trecho do grafo com alguns dos relacionamentos de Lionel Messi:

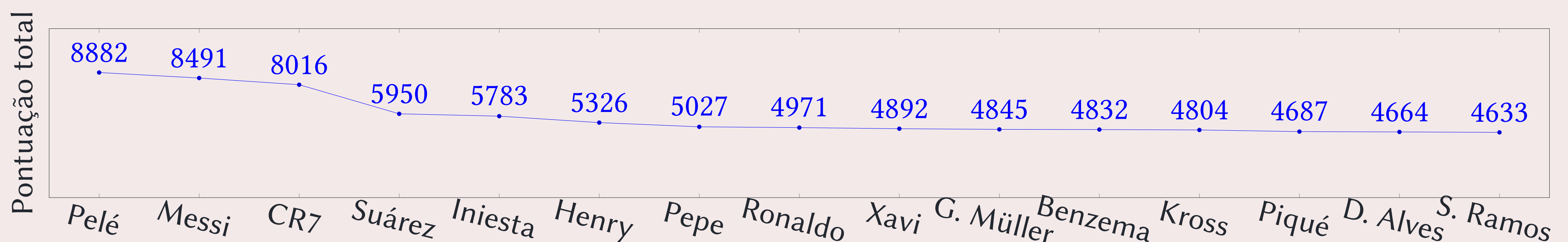


Métricas

- ▶ Uma das diversas métricas possíveis de serem obtidas com a estrutura do grafo é o maior menor caminho entre dois jogadores.
- ▶ Consulta para obter essa métrica no Neo4j:

```
CALL gds.allShortestPaths.stream('graph')
YIELD sourceNodeId, targetNodeId, distance
WITH sourceNodeId, targetNodeId, distance
WHERE isFinite(distance) = true
WITH asNode(sourceNodeId) AS source,
asNode(targetNodeId) AS target, distance
WHERE source <> target
RETURN source, target, distance
```
- ▶ Quando o algoritmo é executado para o grafo com os 768 jogadores, o maior menor caminho encontrado contém 11 nós (10 arestas de distância).

Pontuações totais - Top 15



Bibliografia

- ▶ Elmasri, Ramez and Shamkant Navathe. *Fundamentals of Database Systems*. 7th ed. Addison-Wesley Publishing Company, 2010.
- ▶ Garcia-Molina, Hector et al. *Database Systems: The Complete Book*. 2nd ed. Prentice Hall Press, 2008.
- ▶ Sadalage, Pramod J. and Martin Fowler. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Addison-Wesley Professional, 2012.
- ▶ Salvador, Emanuel Pericles et al. "ModK: Formula for Determining the Best Season and Career of a Football Player by Objective Indicators". In: *The Open Sports Sciences Journal* (2022).

