

Análise e Predição de Química entre Jogadores de Futebol com Base em Dados de Evento

Gabriel Castelo, Matheus Vaz, Victor Kenji, Vinicius Gomes

Universidade Federal de Minas Gerais, Belo Horizonte, Brasil
{gabriel-castelo, vazleal, victorkpa, vinisilvag}@ufmg.br

1 Introdução

No futebol moderno, a avaliação de desempenho ultrapassa os limites da análise estatística individual de cada jogador. Embora métricas como número de passes certos, gols marcados ou desarmes efetuados sejam fundamentais, elas não capturam totalmente a complexidade do jogo coletivo. Um aspecto muitas vezes negligenciado, mas de extrema relevância, é a interação entre os atletas – o que denominamos *química*. Essa conexão e tática entre os jogadores pode influenciar significativamente o rendimento da equipe como um todo. A forma como atletas se comunicam, se posicionam e se antecipam às jogadas uns dos outros é fruto de um entrosamento que vai além das habilidades técnicas individuais.

Apesar de sua importância, a dimensão coletiva do jogo ainda é frequentemente negligenciada em processos de recrutamento e formação de elencos. Um exemplo marcante dessa falha ocorreu na janela de transferências de inverno da temporada 2010/2011, quando o Liverpool contratou simultaneamente Andy Carroll e Luis Suárez [1]. Carroll, um centroavante de referência com forte presença aérea, teve dificuldades de adaptação ao estilo da equipe, resultando em apenas 6 gols marcados em 44 partidas. Em contraste, Suárez – um atacante mais móvel e com características mais alinhadas ao perfil do time – alcançou 69 gols em 110 jogos. Esse contraste evidencia como a falta de química entre jogador e equipe pode comprometer o desempenho, mesmo quando se trata de atletas tecnicamente qualificados.

Nesse sentido, este trabalho propõe investigar os aspectos que influenciam a qualidade do entrosamento entre jogadores de futebol, buscando compreender em que medida atletas conseguem se complementar em campo. O foco está na análise de duplas de jogadores que atuaram juntos, com o objetivo de classificar o nível de química observado entre eles a partir de dados de desempenho coletivo e individual. A partir dessa classificação, queremos também treinar um modelo preditivo capaz de estimar o grau de entrosamento potencial entre jogadores que nunca atuaram juntos.

2 Motivação e Justificativa

A crescente utilização de dados no futebol tem transformado as decisões tomadas pelos grandes times, tanto intra quanto extra campo. Seja em termos de

performance, contratações ou táticas de jogo, um grande número de equipes possui um setor voltado para a geração de estatísticas úteis. No entanto, a maioria das análises está centrada em indicadores e jogadores de forma individual, negligenciando aspectos coletivos, como a sinergia entre os atletas.

Alguns estudos, porém, têm buscado uma abordagem que se estende para além da análise individual, incorporando perspectivas coletivas ao considerar a efetividade de grupos de jogadores na construção de jogadas de impacto. Essas investigações evidenciam um movimento crescente dentro da comunidade científica e da indústria esportiva em direção à valorização de aspectos coletivos do desempenho, como o entrosamento, a coordenação tática e a complementaridade de funções em campo.

Esse interesse tem se intensificado especialmente com a crescente profissionalização e estruturação dos departamentos de *sport analytics* nos clubes de futebol, que agora contam com equipes multidisciplinares e acesso a dados cada vez mais granulares. Nesse cenário, torna-se cada vez mais relevante e pertinente estudar métricas que capturem a dinâmica coletiva das equipes, não apenas para a avaliação de desempenho durante as competições, mas também como critério estratégico na contratação de novos jogadores e na montagem de elencos mais coesos e eficientes. A tendência é que análises voltadas à dimensão coletiva do jogo ganhem espaço e se consolidem como parte fundamental da tomada de decisão no futebol profissional.

3 Problema a investigar

Nesse contexto, o principal problema investigado neste trabalho é: quão bem um time de futebol consegue se entrosar? Mais especificamente, busca-se desenvolver uma forma de quantificar esse entrosamento – ou química – e analisá-lo de maneira sistemática, tanto no aspecto ofensivo quanto defensivo, com base em dados históricos de partidas. O objetivo é construir métricas que permitam avaliar o nível de integração entre os jogadores, indo além da performance individual para capturar elementos do jogo coletivo.

Além disso, este trabalho propõe averiguar a viabilidade de treinar um modelo preditivo capaz de estimar o entrosamento entre pares de jogadores que nunca atuaram juntos. A ideia é avaliar se, à partir de dados prévios sobre estilo de jogo, características físicas e histórico de desempenho, é possível antecipar a compatibilidade entre dois atletas. Caso bem-sucedido, esse modelo poderia ser utilizado como ferramenta auxiliar em processos de montagem de elencos, oferecendo previsões sobre o potencial de química entre jogadores antes mesmo de sua atuação conjunta em campo.

4 Desafios

Para a realização da pesquisa, enfretaremos alguns desafios relevantes. São eles:

- Coleta de indicadores de performance física dos jogadores;

- Coleta de dados que contenham o estilo de jogo dos jogadores analisados;
- A reinterpretação da métrica VAEF e sua adaptação para interações entre dois jogadores, uma vez que queremos usar essa métrica como base para medir a química dos jogadores a partir do impacto de ações conjuntas;
- A escassez de dados posicionais de alta resolução em conjuntos abertos, especialmente para estimar o impacto defensivo das ações;
- O treinamento de um modelo que prevê a química entre jogadores que nunca atuaram juntos, considerando fatores culturais, físicos e táticos.

5 Dados utilizados

Para o cálculo das métricas e o treinamento dos modelos de química ofensiva e defensiva, serão utilizados dados públicos disponibilizados por dois dos principais provedores de estatísticas no futebol: Wyscout e StatsBomb. Ambos são amplamente reconhecidos e adotados tanto na comunidade acadêmica quanto por departamentos de *sport analytics* de clubes profissionais.

Os conjuntos de dados utilizados descrevem, em nível detalhado, as ações e eventos realizados pelos jogadores com a bola ao longo das partidas – os chamados dados de evento. Informações adicionais sobre os atletas, como posição, nacionalidade, idade e altura, também estarão disponíveis. Quando necessário, os dados serão simplificados para facilitar a modelagem, além de serem enriquecidos com variáveis complementares, como perfis de estilo de jogo e indicadores de desempenho físico, a fim de capturar com maior precisão os fatores que influenciam o entrosamento entre jogadores.

6 Cronograma

O cronograma do projeto está dividido em etapas semanais:

- **Semanas 1 – 2:** Processamento dos dados do Wyscout e StatsBomb e estruturação das interações entre jogadores;
- **Semanas 3 – 4:** Implementação do cálculo das métricas de química ofensiva (JOI90) e defensiva (JDI90);
- **Semana 5 – 6:** Treinamento dos modelos de *machine learning* para predição de química entre jogadores que nunca atuaram em conjunto;
- **Semana 7 – 8:** Finalização do relatório e publicação do repositório com código aberto.

Referências

1. Bransen, L., Haaren, J.V.: Player chemistry: Striving for a perfectly balanced soccer team. CoRR **abs/2003.01712** (2020), <https://arxiv.org/abs/2003.01712>