

# Coleção especial de protetor solar - Edição WSL (Circuito de Surf)

## Trabalho 3 de Aprendizado Não-Supervisionado de Máquina

Malcolm dos Reis Alves Pereira - 187642

### Introdução

A temporada do Championship Tour (CT) de 2024 promete ser emocionante, com uma forte presença de surfistas brasileiros se destacando nas competições. O CT é o principal circuito mundial de surfe, organizado pela World Surf League (WSL), reunindo os melhores surfistas do mundo para competir em diversas etapas ao longo do ano. Este campeonato é composto por uma série de eventos realizados em locais icônicos de surfe, como Pipeline no Havaí, Teahupo'o no Taiti e J-Bay na África do Sul, onde os atletas enfrentam ondas desafiadoras e imprevisíveis.

Desse modo, o Brasil, conhecido como o “Brazilian Storm” no mundo do surfe, tem mostrado um desempenho excepcional ao longo dos anos, e 2024 não será diferente. Com atletas de renome mundial, como Gabriel Medina, Italo Ferreira e Tatiana Weston-Webb, os surfistas brasileiros estão prontos para enfrentar os melhores do mundo nas ondas mais desafiadoras do circuito. A presença significativa desses atletas reforça a importância do Brasil no cenário global do surfe e aumenta a expectativa para resultados impressionantes nesta temporada.

Entretanto, exposição ao sol sem a devida proteção pode causar sérios danos à pele, especialmente para surfistas que passam longos períodos na água salgada. A combinação de radiação solar intensa e a reflexão da luz na superfície do mar aumenta o risco de queimaduras solares, envelhecimento precoce e, em casos mais graves, câncer de pele. Portanto, é essencial que cada atleta utilize um protetor solar adequado ao seu tom de pele, garantindo uma proteção eficaz sem a necessidade de aplicar o produto em excesso. Protetores solares personalizados, desenvolvidos com fatores de proteção específicos para cada tipo de pele, são cruciais para manter a saúde dos surfistas, permitindo que eles se concentrem em suas performances com segurança e confiança.

Com isso, suponha que uma empresa brasileira de cosméticos *BrasilisCosméticos* irá patrocinar esses atletas e, por essa razão, precisa produzir protetores solares de acordo com a necessidade de cada grupo de pessoas com o mesmo tom de pele. Portanto, será feita uma clusterização entre os surfistas brasileiros que participarão desse campeonato através do tom de pele, identificando esses grupos e possibilitando determinar qual fator de proteção cada grupo precisará para que os atletas estejam seguros contra os raios solares durante a competição.

### Materiais e Métodos

Este estudo utilizou técnicas de aprendizado não supervisionado para clusterizar rostos de surfistas brasileiros com base no tom de pele, visando a criação de protetores solares personalizados. As etapas incluem a identificação dos rostos, a conversão das imagens para uma matriz de duas dimensões sem cor, e a clusterização dos dados.

As imagens dos surfistas foram obtidas e carregadas para processamento. Para identificar os rostos nas fotos, utilizamos técnicas de detecção facial para isolar os rostos e gerar máscaras correspondentes. As imagens foram então convertidas para escala de cinza para focar apenas na intensidade dos pixels, removendo a



Figure 1: Fotos dos surfistas que irão participar do campeonato

informação de cor. Essas imagens em escala de cinza foram uniformemente redimensionadas para garantir a consistência das dimensões.



Figure 2: Identificação dos rostos e colocando a imagem em apenas duas dimensões (sem cores)

As imagens em escala de cinza foram transformadas em vetores numéricos, onde cada pixel foi representado por um valor. Estes vetores foram utilizados para calcular a distância entre as imagens, formando a base para a clusterização hierárquica. A clusterização hierárquica foi realizada para agrupar as imagens em clusters baseados na similaridade de tons de pele.

Para determinar o número ideal de clusters, aplicamos o método do cotovelo (Elbow Method). Este método envolve a execução de uma série de algoritmos de k-means com diferentes números de clusters e a plotagem da variância total explicada por cada configuração. A escolha do número de clusters é baseada no ponto onde o aumento da variância total começa a diminuir significativamente, indicando um bom balanço entre a simplicidade do modelo e a explicação da variância dos dados.

Para entender melhor a estrutura dos dados e visualizar a distribuição dos clusters, realizamos uma Análise de Componentes Principais (PCA). A PCA foi utilizada para reduzir a dimensionalidade dos dados, transformando as variáveis originais (pixels das imagens) em um conjunto menor de variáveis não correlacionadas, chamadas componentes principais.

Com 9 componentes principais dispostas, pôde ser observado que as duas primeiras componentes já são suficientes para explicar a variabilidade dos dados. Portanto, na visualização dos grupos, será elas que vamos utilizar.

## Resultados

Começando pelo método de Elbow, a partir do gráfico de pontos da Figura 3 que mostra a variância considerando o número de grupos existentes, é possível observar uma leve inflexão quando temos  $K = 3$  grupos, como indicado pela linha vermelha. Essa leve inflexão foi considerada suficiente para a determinação do número de clusters que iremos trabalhar.

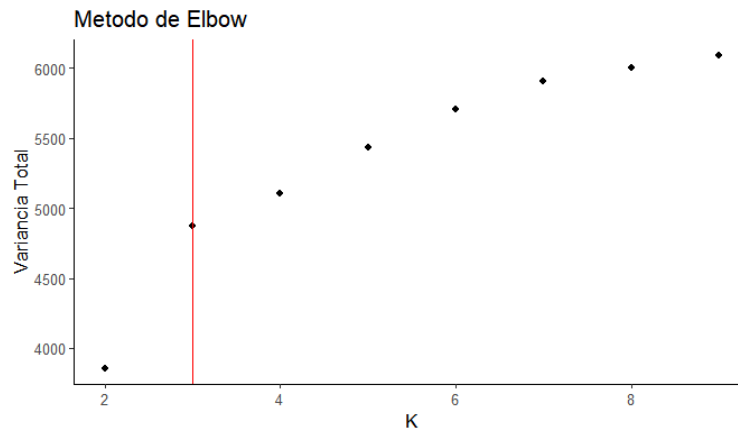


Figure 3: Andamento da variância por número de cluster (Método de Elbow)

Com o número de grupos ideal já obtido, conseguimos o  $K = 3$ , com isso, pôde ser feito o corte no dendrograma da hierarquia gerada pelo método *average linkage* em cima desses dados, que está indicado pela linha vermelha no gráfico do Dendrograma do cluster da Figura 4.

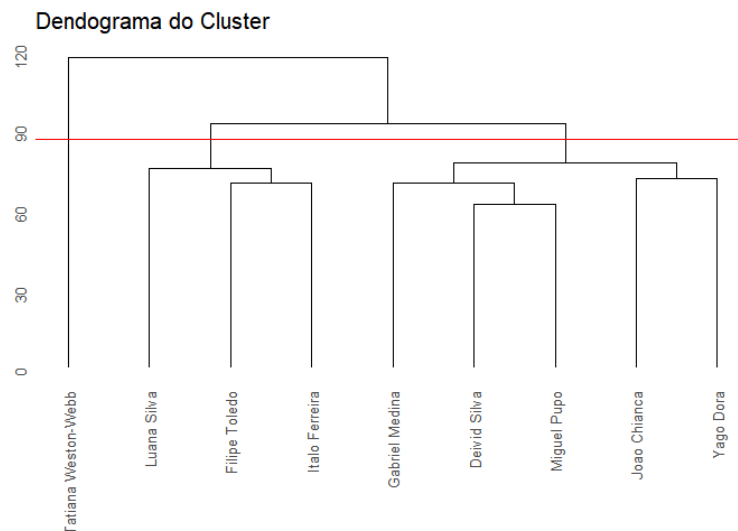


Figure 4: Dendrograma exibindo os grupos formados e o corte na árvore

A partir dessas considerações, temos cada cluster bem definido, como é mostrado no Diagrama de pontos que coloca os pontos em seus respectivos grupos da Figura 5. Exemplificando, na Figura 6 tem um surfista de cada grupo que está sendo mostrado na Figura 5.

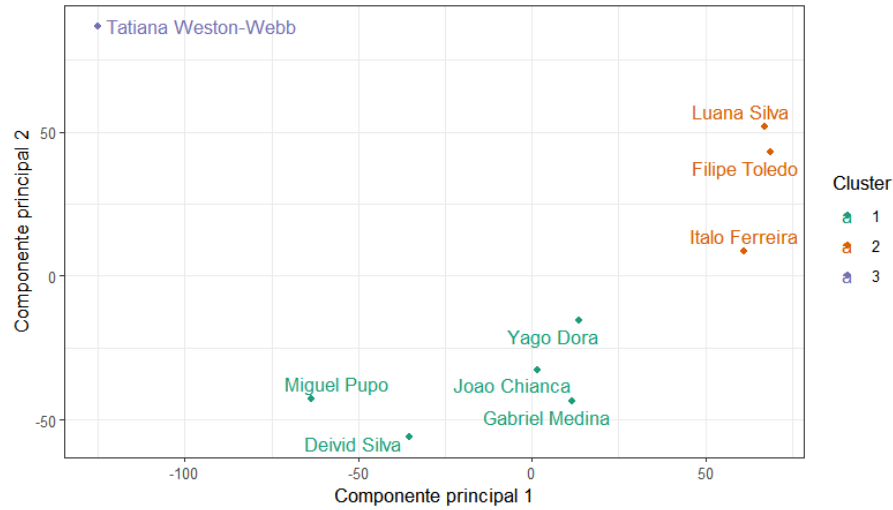


Figure 5: Grupos identificados através do clustering hierárquico

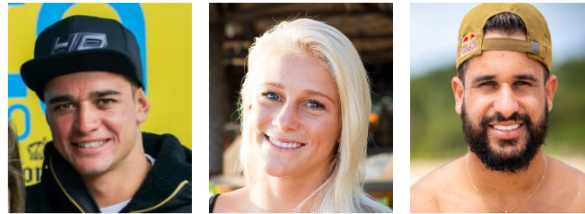


Figure 6: Deivid Silva, Tatiana e Italo Ferreira, respectivamente

## Discussão

Agora, observando os grupos de atletas que possuem similaridades em seus tons de pele e, conseqüentemente, em suas necessidades nas configurações dos protetores solares, cabe à BrasilisCosméticos determinar o fator de proteção que cada grupo precisa. Com esta análise detalhada, a empresa pode desenvolver produtos específicos que ofereçam a proteção solar ideal para cada grupo, evitando a aplicação excessiva de protetor solar e garantindo que os surfistas estejam bem protegidos contra os raios solares durante suas competições. Esta abordagem personalizada não só atende às necessidades individuais dos atletas, mas também maximiza a eficácia dos produtos, promovendo a saúde e o desempenho dos surfistas brasileiros.