

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

Luiz Antônio Nonenmacher Júnior

APLICAÇÃO DE TÉCNICAS MULTIVARIADAS EM
PAINÉIS SENSORIAIS DE CAFÉ

Porto Alegre

2018

Luiz Antônio Nonenmacher Júnior

Aplicação de técnicas multivariadas em painéis sensoriais de café

Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio Grande do Sul como requisito parcial à obtenção do título de Mestre em Engenharia de Produção, modalidade Acadêmica, na área de concentração em Sistemas de Qualidade.

Orientador: Michel José Anzanello,
Ph.D.

Porto Alegre

2018

Luiz Antônio Nonenmacher Júnior

Aplicação de técnicas multivariadas em painéis sensoriais de café

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia de Produção na modalidade Acadêmica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora designada pelo Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Rio Grande do Sul.

—
Prof. Michel José Anzanello, *Ph.D.*
Orientador PPGEP/UFRGS

—
Prof. Flávio Sanson Fogliatto, *Ph.D.*
Coordenador PPGEP/UFRGS

Banca Examinadora:

Profa. Carla ten Caten, Dra. (PPGEP/UFRGS)

Prof. Lucas Louzada Pereira, Dr. (IFES)

Alessandro Kahmann, Dr. (DEPROT)

NONENMACHER JUNIOR, L. A. *Aplicação de técnicas multivariadas em painéis sensoriais de café*, 2018. Dissertação (Mestrado em Engenharia) – Universidade Federal do Rio Grande do Sul, Brasil.

RESUMO

No setor alimentício é frequente o uso de painéis sensoriais. Esses painéis são formados por um grupo de degustadores treinados que avaliam amostras de produtos em diversos atributos sensoriais e podem ser utilizados para garantir a autenticidade de produtos, assegurar sua qualidade e obter suas características sensoriais. Esta dissertação visa a aplicar técnicas multivariadas em painéis sensoriais de café, com o objetivo de calibrar e selecionar degustadores, prever notas conferidas pelos mesmos e classificar produtos em categorias de qualidade. Para atingir esses objetivos, inicialmente utiliza-se o Índice Alfa de Ledauphin e um índice de aprendizado para simular o efeito do treinamento de degustadores e comparar o desempenho de três diferentes métodos de treinamento. Na sequência, aplica-se uma regressão por mínimos quadrados parciais (PLSR) para prever as notas conferidas pelos degustadores com base em variáveis físico-químicas, selecionando as variáveis mais relevantes para predição através de um índice de importância de variáveis. Por fim, compara-se o desempenho combinado de dois índices de seleção de variáveis baseados em distância de Bhattacharyya e análise de componentes principais (PCA), e três técnicas de classificação - algoritmo do vizinho mais próximo (k -NN), análise linear discriminante (LDA) e rede neural probabilística (PNN) – na classificação de amostras em categorias de qualidade com base em dados espectrais. Com a aplicação dessas técnicas foi possível aprimorar a capacidade preditiva e reduzir a quantidade de variáveis e avaliadores necessários, reduzindo o custo de futuros experimentos.

Palavras-chave: Painéis sensoriais, Café, Técnicas multivariadas, Seleção de variáveis, Índice Alfa de Ledauphin.

NONENMACHER JUNIOR, L. A. *Application of multivariate techniques in coffee sensory panels*. 2018. Dissertation (Master in Engineering) - Federal University of do Rio Grande do Sul, Brazil.

ABSTRACT

In the food industry, the utilization of sensory panels is common. Those panels are composed of a group of trained graders that evaluate some samples in a group of sensory attributes, and can be utilized to guarantee the product's authenticity, assure its quality and obtain its sensory characteristics. This dissertation aims to apply multivariate techniques in sensory panels of coffee products, to train and select graders, forecast the grades given by them and classify products in quality categories. For those purposes, initially the Alfa Index of Ledauphin and a learning index are utilized to simulate the effect of training and compare the performance of three methods of evaluators training. After, a partial least squares regression (PLSR) is applied to forecast the grades given by graders based on physical and chemical variables of the samples, selecting the variables with the best performance using a variable importance index. Last, the combined performance of two variable importance indexes, based on the Bhattacharyya distance and the principal component analysis (PCA), and three classification techniques – the k-nearest neighbor algorithm (k-NN), the linear discriminant analysis (LDA) and the probabilistic neural network (PNN) – are compared when classifying samples based on spectral data. With the use of those techniques, it was possible to increase the predictive capacity and reduce the amount of variables and evaluators utilized, reducing the cost of future experiments.

Keywords: *Sensory panels, Coffee, Multivariate Techniques, Variable Selection, Index Alfa of Ledauphin.*

LISTA DE FIGURAS

Figura 2.1: Comparação dos métodos com $it = 1$	32
Figura 2.2: Comparação dos métodos com $it = 0,75$	32
Figura 2.3: Comparação dos índices com $it = 0,5$	32
Figura 2.4: Comparação dos índices com $it = 0,25$	33
Figura 3.1: Componentes retidos.....	48
Figura 3.2: Seleção de degustadores	49
Figura 3.3: Seleção de variáveis	50
Figura 3.4: Comparação com enumeração total	51

LISTA DE TABELAS

Tabela 2.1: Índices Iniciais.....	28
Tabela 2.2: Treinamento dos atributos	29
Tabela 2.3: Treinamento dos degustadores	30
Tabela 2.4: Treinamento combinado - parcial.....	31
Tabela 2.5: Índices Alfa.....	33
Tabela 2.6: Média de substituições necessárias	34
Tabela 3.1: Alfa de Ledauphin	49
Tabela 3.2: Índice de seleção de variáveis	50
Tabela 4.1: Acurácia e % de variáveis retidas para os 2 índices de importância	67
Tabela 4.2: Acurácia e % de variáveis retidas para as 3 técnicas de classificação	68
Tabela 4.3: Coeficiente de variação de cada método	69
Tabela 4.4: Dados por atributo	70

SUMÁRIO

1. Introdução	11
1.1 Considerações Iniciais	11
1.2 Objetivos.....	12
1.3 Justificativa do Tema e dos Objetivos	12
1.4 Procedimentos Metodológicos.....	13
1.5 Estrutura da Dissertação	14
1.6 Delimitações do Estudo	15
1.7 Referências	16
2. Primeiro artigo: Proposição de alternativas de treinamento de degustadores em um painel sensorial utilizando índices de desempenho	18
2.1. Introdução.....	19
2.2 Referencial teórico.....	20
2.2.1 Abordagens para avaliação de desempenho de painéis sensoriais ..	20
2.2.2. Alfa de Ledauphin	23
2.3. Procedimentos metodológicos	24
2.4. Resultados.....	27
2.4.1 Cálculo do Alfa de Ledauphin inicial.....	28
2.4.2 Treinamento dos atributos	29
2.4.3 Treinamento dos degustadores	29
2.4.4 Treinamento combinado.....	30
2.4.5 Comparação dos resultados	31

2.5. Conclusões.....	34
2.6. Referências	35
3. Segundo artigo: Previsão de notas de atributos em um painel sensorial de café apoiada em PLS e técnicas de seleção.....	38
3.1. Introdução.....	39
3.2. Referencial teórico.....	40
3.2.1 Técnicas de regressão.....	40
3.2.2 Avaliação de erro de predição	42
3.2.3 Seleção de variáveis	44
3.2.4 Seleção de degustadores.....	44
3.3. Procedimentos metodológicos.....	46
3.4. Resultados.....	47
3.4.1 Definição do número de componentes	48
3.4.2 Seleção de degustadores.....	48
3.4.3 Seleção de variáveis	49
3.5. Conclusões.....	52
3.6. Referências	53
4. Terceiro artigo: Comparação de técnicas de classificação e seleção de variáveis aplicadas a amostras de café descritas por variáveis espectrais	56
4.1. Introdução.....	57
4.2. Referencial teórico.....	58
4.2.1 Técnicas de classificação.....	59

	10
4.2.2 Algoritmo do vizinho mais próximo	60
4.2.3 Análise discriminante linear	61
4.2.4 Redes neurais probabilísticas	62
4.2.5 Seleção de variáveis	62
4.3. Procedimentos metodológicos	64
4.4. Resultados.....	66
4.4.1 Análise dos índices	66
4.4.2 Análise das técnicas de classificação	67
4.4.3 Análise dos atributos	70
4.5. Conclusões.....	70
4.6. Referências	71
5. Considerações finais	75
5.1 Conclusões.....	75
5.2 Sugestões para trabalhos futuros	76

1. Introdução

Esta seção apresenta a dissertação, iniciando pela introdução do tema e seguindo pelos objetivos e justificativa dos mesmos. Na sequência apresentam-se os procedimentos adotados para atingimento dos objetivos e finaliza-se a seção com a estrutura da dissertação e as delimitações do estudo.

1.1 Considerações Iniciais

Técnicas multivariadas são utilizadas em situações nas quais um conjunto de mensurações é coletado para cada amostra ou indivíduo, podendo ser aplicadas com objetivo descritivo ou inferencial (RENCHER, 2002). O uso dessas técnicas é crescente nos mais diversos setores, inclusive na indústria de alimentos (CALLAO e RUISÁNCHEZ, 2018).

A indústria alimentícia é um dos setores mais importantes da economia brasileira, representando, em 2017, 9,8% do PIB e 17,9% da exportação do país (ABIA, 2018). Um dos produtos de destaque deste setor é o café, segunda bebida mais consumida no mundo, cujo consumo brasileiro em 2017 foi de 1,07 milhões de toneladas, tornando o país o segundo maior consumidor de café do mundo (EMBRAPA, 2018).

No setor alimentício é muito frequente a utilização de painéis sensoriais. O objetivo principal de um painel é a condução de testes válidos e confiáveis que gerem dados para a tomada de decisão, podendo esses dados ser utilizados para assegurar a qualidade de um produto, garantir sua autenticidade, definir seu preço ou obter suas características sensoriais (MEILGAARD et al., 2007). Um painel é composto por um grupo de degustadores treinados que avaliam amostras de produtos, conferindo a cada amostra uma nota para cada atributo sensorial avaliado (MOSER et al., 2018).

Nesta dissertação serão utilizados dados provenientes de painéis sensoriais de cafés especiais, cujas amostras foram processadas por via-úmida e são provenientes da região Sudoeste Serrana do Espírito Santo. Os painéis sensoriais foram compostos por degustadores treinados e certificados pelo *Coffe Quality Institute (CQI)* e seguiram o protocolo da *Specialty Coffe Association of America (SCA)*. Nesse protocolo estão descritos os procedimentos para preparação das amostras e avaliação de dez atributos

sensoriais (fragrância / aroma, uniformidade, ausência de defeitos, doçura, sabor, acidez, corpo, finalização, equilíbrio e final) (SCA, 2018).

1.2 Objetivos

O objeto principal da dissertação é aplicar técnicas multivariadas sobre dados de painéis sensoriais de café com vistas a mensurar o desempenho de degustadores e prever notas/classes das amostras. Esse objetivo pode ser desdobrado nos seguintes objetivos específicos:

(i) Comparar o aprimoramento de desempenho esperado pela aplicação de três métodos de treinamento a degustadores de um painel sensorial, definindo o método que se destaca em cada situação;

(ii) Aplicar a regressão PLS para prever notas conferidas por degustadores a amostras de café em relação a dez atributos sensoriais, utilizando variáveis de processo e físico-químicas como variáveis independentes do modelo;

(iii) Utilizar o Índice Alfa de Ledauphin e um índice baseado em PLS para selecionar degustadores e variáveis a fim de aumentar a capacidade preditiva da regressão PLS; e

(iv) Comparar o desempenho de dois índices de seleção de variáveis (baseados em PCA e distância de Bhattacharyya) e três técnicas de classificação (k -NN, LDA e PNN) utilizadas para a classificação de amostras de café em duas categorias de qualidade.

1.3 Justificativa do Tema e dos Objetivos

Na literatura acadêmica encontram-se aplicações de painéis sensoriais aliados a técnicas quantitativas para diversos fins, dentre os quais destacam-se atestar a origem geográfica das amostras (PÉREZ-ELORTONDO et al., 2018), discriminar produtos (OBERRAUTTER, 2018), definir o número ótimo de degustadores de um painel (TALSMA, 2018) e avaliar características desejadas por consumidores (SASAKI et al., 2014). Apesar das diversas abordagens voltadas ao cumprimento dos objetivos acima, entende-se como relevante a possibilidade de desenvolvimento de técnicas mais robustas em termos de predição das variáveis sensoriais a partir de variáveis laboratoriais descrevendo as amostras, bem como para seleção dos degustadores mais

consistentes. Tais demandas teóricas justificam a presente dissertação em termos acadêmicos.

Sob perspectiva prática, a aplicação das técnicas multivariadas tem dois objetivos de interesse para produtores de café: redução de custo e aumento da qualidade do produto. A redução do custo se manifesta na seleção de degustadores, o que diminui despesas relacionadas à execução de futuros painéis sensoriais. Além disso, ao encontrar a melhor metodologia de treinamento dentre as três propostas, reduz-se o custo com treinamento. Por fim, a seleção de variáveis também diminui os custos de futuros experimentos, reduzindo despesas associadas a procedimentos laboratoriais. O aumento da qualidade está relacionado tanto à previsão das notas conferidas pelos degustadores quanto à classificação das amostras em categorias de qualidade, pois elas fornecem informação que permitem aos produtores identificar quais fatores de seu processo produtivo tem mais efeito na qualidade do café. De tal forma, tais objetivos justificam o presente trabalho no âmbito prático.

1.4 Procedimentos Metodológicos

A presente pesquisa é de natureza aplicada, pois utiliza a base teórica para a solução de problemas práticos, e tem uma abordagem quantitativa, pelo uso de modelagem matemática e técnicas estatísticas para solução dos problemas. Quanto aos procedimentos técnicos, a pesquisa se classifica como *Expost-facto*, pois se utilizam os dados após ocorrência dos fatos, não havendo controle das variáveis independentes. Quanto aos objetivos, é tida como explicativa porque visa a identificar as variáveis que contribuem para a ocorrência dos fenômenos (DA SILVA e MENEZES, 2005).

O método proposto pelo primeiro artigo inicia-se pelo cálculo do Alfa de Ledauphin do painel, obtido através das médias dos índices Alfa de cada degustador, e pelo cálculo da matriz ponderada de notas. Através dessa matriz ponderada e de um índice de aprendizado simula-se o efeito do treinamento de um degustador. Na sequência são simulados os efeitos dos três tipos de treinamento: treinamento de atributos para todos os degustadores, treinamento de degustadores em todos os atributos e treinamento de cada atributo para cada degustador, sendo cada treinamento simulado com quatro valores de índice de aprendizado. Por fim, os tipos de treinamento são comparados para cada valor de índice de aprendizado e define-se o treinamento para cada situação.

Para prever as notas conferidas por degustadores às amostras de café, o segundo artigo é executado em três etapas. Na primeira etapa aplica-se a regressão PLS, definindo-se o número de componentes a ser utilizado com base no erro de predição. A partir disso, a segunda etapa tem por objetivo selecionar os degustadores que maximizem a capacidade preditiva do modelo. Para isto, os degustadores são ordenados pelo Índice Alfa de Ledaphin e, a cada rodada, retira-se o degustador de menor índice e se recalcula o erro de predição. Ao fim deste processo, seleciona-se o número de degustadores que conduz ao menor erro. Por fim, a terceira etapa faz a seleção de variáveis (físico-químicas e de processo) utilizando um processo semelhante ao da segunda etapa, com a utilização de um índice de importância baseado em parâmetros do PLS ao invés do Alfa de Ledaphin.

O terceiro artigo tem como objetivo comparar o desempenho combinado de dois índices de seleção de variáveis e três técnicas de classificação para a classificação de amostras de café em categorias de qualidade com base em dados espectrais. O primeiro passo consiste em processar as variáveis espectrais pelo método SNV (*Standard normal variate*), corrigindo os efeitos de dispersão da luz. Após isso, calculam-se dois índices de importância para cada variável, baseados em coeficientes do PCA e distância de Bhattacharyya. Na sequência, são executados os experimentos para cada uma das 42 combinações de técnicas (*k*-NN, LDA e PNN), índices e atributos, sendo que cada experimento consiste em aplicar a técnica, calcular a acurácia de classificação, retirar o degustador de pior índice e calcular novamente a acurácia, até restar apenas uma variável. Por fim, os índices e as técnicas de classificação são comparados entre si a fim de encontrar a melhor combinação; faz-se então uma análise de cada atributo.

1.5 Estrutura da Dissertação

A dissertação é dividida em cinco capítulos. O primeiro capítulo introduz o tema, seguido pela apresentação dos objetivos e justificativa do trabalho. Este capítulo é finalizado com a apresentação dos procedimentos metodológicos e delimitações da pesquisa.

O segundo capítulo é composto pelo primeiro artigo, que compara três métodos de treinamento de degustadores (treinamento de atributos, treinamento de degustadores e treinamento combinado) de um painel sensorial de café. O treinamento é simulado

através de um índice de aprendizado e de uma matriz de notas ponderadas e o desempenho do treinamento é obtido pela utilização do Índice Alfa de Ledauphin.

O terceiro capítulo apresenta o segundo artigo, que visa a aplicar a regressão PLS para prever notas conferidas por degustadores a amostras de café, utilizando variáveis de processo e físico-químicas como variáveis independentes. Juntamente à regressão é realizada uma seleção de degustadores e variáveis que tem como objetivo aumentar a capacidade preditiva do modelo e reduzir o custo de futuros experimentos. A seleção de degustadores é executada com o uso do Índice Alfa de Ledauphin e da seleção de variáveis com um índice de importância baseado em parâmetros do PLS.

No quarto capítulo vem o terceiro artigo, cujo objetivo é comparar o desempenho de combinações de técnicas de classificação e índices de importância de variáveis na classificação de amostras de café em faixas de qualidade com base em variáveis espectrais. São comparados dois índices (um baseado em parâmetros do PCA e outro na distância de Bhattacharyya), e três técnicas de classificação (k -NN, LDA e PNN); as combinações de técnicas e índices são aplicadas para sete atributos sensoriais do produto.

Por fim, o quinto capítulo apresenta a conclusão do trabalho, avaliando os resultados da dissertação com base nos objetivos e delimitações da mesma. Além disso, esse capítulo apresenta sugestões para pesquisas e desdobramentos futuros no tema.

1.6 Delimitações do Estudo

Este estudo apresenta as seguintes delimitações:

- As técnicas são aplicadas apenas em dados provenientes de painéis sensoriais de amostras de café de uma região específica do Brasil, não sendo os resultados replicados para outras amostras de café ou diferentes produtos alimentícios;
- Não são propostas novas ferramentas de previsão, classificação e seleção de variáveis, sendo apenas utilizadas combinações dessas técnicas para atingimento dos objetivos;

- As seleções de variáveis e degustadores utilizam somente técnicas baseadas em índices de importância, não sendo utilizadas outras abordagens de seleção; e
- O treinamento de degustadores não é aplicado em situação prática, sendo seus efeitos simulados através de um índice de aprendizado e de uma matriz ponderada de notas.

1.7 Referências

ABIA. **Faturamento 2017**. 2018. Disponível em <<https://www.abia.org.br/vsn/anexos/faturamento2017.pdf>> Acesso em 25 jun. 2018.

ABIA. **Balança comercial 2017**. 2018. Disponível em <<https://www.abia.org.br/vsn/anexos/balancacomercial2017.pdf>> Acesso em 25 jun. 2018.

CALLAO, M.; RUISÁNCHEZ, I. An overview of multivariate qualitative methods for fraud detection. **Food Control**, Volume 86, 2018.

DA SILVA, E.; MENEZES, E. **Metodologia de Pesquisa e Elaboração de Dissertação**. 4. Ed., Florianópolis: UFSC, 2005.

EMBRAPA. **Consumo café**. 2018. Disponível em <<https://www.embrapa.br/busca-de-noticias/-/noticia/31768082/brasil-consome-215-milhoes-de-sacas-de-cafe-em-2017>> Acesso em 25 jun. 2018.

LEDAUPHIN, S.; HANAFI, M.; QANNARI, E. Assessment of the agreement among the subjects in fixed vocabulary profiling. **Food Quality and Preference**, Volume 17, 2006.

MEILGAARD, M.; CIVILLE, G.; CARR, B. **Sensory Evaluation Techniques**. 4 Ed., Boca Raton: CRC Press, 2007.

MOSER, M.; LEPAGE, M.; PINEAU, N.; FILLION, L.; RYTZ, A. Replicates in sensory profiling: Quantification of the impact of moving from two to one assessments. **Food Quality and Preference**, Volume 65, 2018.

OBERRAUTER, L.; JANUSZEWSKA, R.; SCHLICH, P.; MAJCHRZAK, D. Sensory evaluation of dark origin and non-origin chocolates applying Temporal Dominance of Sensations (TDS). **Food Research International**, Volume 111, 2018.

PÉREZ-ELORTONDO, F.; SYMONEAUX, R.; ETAIO, I.; COULON-LEROY, C.; MAÎTRE, I.; ZANNONI, M. Current status and perspectives of the official sensory control methods in protected designation of origin food products and wines. **Food control**, Volume 88, 2018.

RENCHE, A. **Methods of Multivariate Analysis**, 2nd Ed., John Wiley & Sons, 2002.

SASAKI, K.; MOTOYAMA, M.; NARITA, T.; HAGI, T.; OJIMA, K.; OE, M.; NAKAJIMA, I.; KITSUNAI, K.; SAITO, Y.; HATORI, H.; MUROYA, S.; NOMURA, M.; MIYAGUCHI, Y.; CHIKUNI, K. Characterization and classification of Japanese consumer perceptions for beef tenderness using descriptive texture characteristics assessed by a trained sensory panel. **Meat Science**, Volume 96, 2014.

SCA, 2018. **Specialty Coffee Association Protocols**. 2018. Disponível em [≤https://sca.coffee/research/protocols-best-practices>](https://sca.coffee/research/protocols-best-practices). Acesso em 4 set. 2018.

TALSMA, P. How much sensory panel data do we need?. **Food quality and preference**, Volume 67, 2018.

2. Primeiro artigo: Proposição de alternativas de treinamento de degustadores em um painel sensorial utilizando índices de desempenho

Luiz Antônio Nonenmacher Júnior (ljuniornone@gmail.com)

Michel José Anzanello (anzanello@producao.ufrgs.br)

Resumo

Painéis sensoriais são muito utilizados na indústria de alimentos para controle de qualidade, classificação de produtos em categorias e pesquisas com diversos fins. Um painel sensorial é composto por um grupo de degustadores que confere notas a diferentes atributos sensoriais de amostras de produtos; em tal processo, espera-se que o painel seja composto por membros que avaliem produtos iguais de forma semelhante e consigam discriminar produtos diferentes. O objetivo deste artigo é aplicar e comparar três abordagens de treinamento de degustadores, com o objetivo de melhorar o desempenho do painel, medido através do Índice Alfa de Ledauphin. As três abordagens comparadas foram (i) treinamento de alguns degustadores para todos os atributos (treinamento de degustadores), (ii) treinamento de alguns atributos para todos os degustadores (treinamento de atributos), e (iii) combinação dessas duas técnicas (alguns atributos para alguns degustadores). O método combinado foi o que apresentou os melhores resultados, com uma necessidade de treinamento, para um mesmo desempenho do painel, 18,7% menor que o treinamento de degustadores e 51,4% menor que o treinamento de atributos.

Palavras-chave: Painéis Sensoriais, Treinamento de Degustadores, Índice Alfa de Ledauphin.

Abstract

Sensory panels are commonly utilized in the food industry for purposes of quality control, classification and research. A sensory panel is composed of a group of graders that evaluate some products samples in a group of attributes. It is expected that the graders which compose a panel are capable of evaluating equally the same products and capable of discriminating different ones. The objective of this article is to apply and compare three methods of training of graders that aim to increase the panel performance, evaluated by the Alpha Index of Ledauphin. The three compared methods

are: (i) training of some graders in all attributes (graders training), (ii) training of some attributes for all graders (attributes training) and (iii) the combination of this techniques (some attributes for some graders). The combined method is the one that presented the best results, with a training need, for the same panel performance, 18.7% lower than the graders training and 51.4% lower than the attributes training.

Keywords: Sensory Panels, Graders Training, Alpha Index of Ledauphin.

2.1. Introdução

Na indústria de alimentos é comum o uso de painéis sensoriais para se avaliar um grupo de características sensoriais de um produto (WILDERJANS e CARIOU, 2016) com vistas ao controle de qualidade e classificação dos produtos em faixas de atribuição de valor (por exemplo, faixas *premium* e regular), dentre outros (BANFIELD e HARRIES, 1975). Um painel é formado por um grupo de degustadores treinados que conferem notas a amostras de produtos (com a presença ou não de repetições) descritos por um conjunto de atributos, sendo os resultados tipicamente expressos numa escala numérica (ANZANELLO et al., 2011).

Para assegurar que um painel tenha um bom desempenho, é fundamental que seus integrantes sejam simultaneamente capazes de avaliar produtos de forma semelhante entre si e de discriminar produtos diferentes. Existem diferentes técnicas para mensurar o desempenho de um painel, sendo que nesse artigo será utilizado o Índice Alfa de Ledauphin (LEDAUPHIN et al., 2006). Tal índice é baseado no autovetor β , obtido a partir do maior autovalor de uma matriz \mathbf{S} gerada a partir da comparação de matrizes que descrevem as notas individuais de cada degustador do painel sensorial. Essa técnica gera dois principais resultados: um índice que mensura a concordância do painel como um todo e uma série de valores que avalia o desempenho de cada um dos degustadores.

Este artigo propõe a comparação de três cursos de ação distintos com vistas ao treinamento de degustadores, monitorando seu resultado através do Alfa de Ledauphin do painel. Um desses métodos, mais utilizado na prática, consiste em selecionar um ou mais degustadores e treiná-lo em todos os atributos. Um dos métodos propostos consiste em selecionar um ou mais atributos e treiná-lo dentre todos os degustadores. Por fim, o último método proposto consiste em conduzir um treinamento de forma combinada, selecionando apenas alguns atributos de alguns degustadores para realizar o

treinamento. O impacto do treinamento sobre as notas de cada degustador será simulado com a utilização de uma matriz ponderada de notas e um índice de aprendizado.

Este artigo está organizado em cinco seções. Após a introdução, a seção 2.2 apresenta algumas técnicas utilizadas para avaliar o desempenho de painéis sensoriais, com enfoque no Índice Alfa de Ledauphin. Já a seção 2.3 apresenta o método proposto para o atingimento dos objetivos deste artigo, enquanto os resultados da aplicação e uma análise dos mesmos compõem a seção 2.4. Por fim, a seção 2.5 conclui o artigo, retomando os objetivos apresentados na introdução e sugerindo pesquisas futuras na área.

2.2 Referencial teórico

Essa seção inicia com a definição de painéis sensoriais e a apresentação de algumas técnicas utilizadas para avaliar o desempenho de tais painéis. Na sequência, as premissas do Índice Alfa de Ledauphin, utilizado neste trabalho, são descritas em maiores detalhes.

2.2.1 Abordagens para avaliação de desempenho de painéis sensoriais

Um painel sensorial que avalia um produto alimentício é composto por um grupo de degustadores treinados que graduam, em geral numa escala numérica, um conjunto de atributos para diferentes amostras de um grupo de produtos. Para que o painel sensorial tenha sucesso em sua implementação, é desejado que os degustadores sejam capazes de (i) repetir consistentemente suas avaliações ao longo do tempo, (ii) avaliar produtos de maneira semelhante aos demais membros do painel, e (iii) discriminar produtos com características sensoriais diferentes (LATREILLE et al., 2006). Existem diferentes metodologias para fazer a avaliação de desempenho de um painel e de seus membros, conforme apresentadas a seguir.

Uma das abordagens básicas para avaliação de desempenho de painéis sensoriais é utilizar a variância dos dados para calcular índices numéricos que têm sua significância testada através de métodos estatísticos como a ANOVA ou o Teste T. Na sequência, os índices podem formar um espaço bidimensional e os degustadores, atributos ou produtos podem ser analisados graficamente. Um conjunto desses índices é utilizado por Rossi (2012) para avaliar degustadores. São utilizados dois indicadores: índice de repetitividade (*repeatability*), que avalia a capacidade do degustador de avaliar repetições do mesmo produto de forma igual, e o índice de reprodutibilidade

(*reproducibility*), que mede a capacidade de um degustador, na média, ter uma mesma avaliação que outros membros do painel. Este é um dos métodos comparados por Pinto et al. (2014) para classificar degustadores de acordo com seu desempenho.

Outro conjunto de índices foi proposto por Bi (2007) para avaliar a qualidade de um painel sensorial. Ele utiliza o coeficiente de concordância e o de confiabilidade (*agreement and reliability*), sendo que o coeficiente de concordância avalia o consenso entre os diferentes degustadores para um mesmo produto e o de confiabilidade mede a capacidade de discriminação do painel. Com propósitos semelhantes, Talsma (2016) utiliza a teoria da generalização para avaliar o desempenho de um painel sensorial e de seus membros. A variância é decomposta em sete elementos (avaliando a variância do produto, do degustador, da réplica e suas interações de segunda e terceira ordem); o indicador de confiabilidade, por sua vez, é calculado através da divisão da variância de cada componente pela variância total.

Uma técnica frequentemente utilizada para avaliação de desempenho de painéis é a análise de componentes principais (PCA). A técnica visa a transformar as variáveis originais em um conjunto reduzido de componentes principais (combinações lineares das variáveis originais) com o objetivo de explicar o máximo possível da variabilidade dos dados originais (BORGOGNONE et al., 2001). O PCA forma a base da análise de consonância (*consonance analysis*), técnica proposta por Dijksterhuis (1995), que avalia o desempenho dos degustadores através da comparação do primeiro autovalor com os demais. A análise de consonância também é um dos métodos utilizados por Pinto et al. (2014) para classificar degustadores de acordo com seu desempenho.

O PCA pode ser utilizado conjuntamente com outras técnicas com vistas à avaliação de degustadores. Anzanello et al. (2011) utilizaram PCA e *k-Nearest Neighbor* (*k*-NN) para encontrar um conjunto ótimo de atributos e degustadores que melhor diferenciam as amostras analisadas, enquanto que Sinesio et al. (1990) usaram PCA, análise generalizada de Procrustes (GPA) e regressão de mínimos quadrados parciais (PLSR) para avaliar a capacidade de um grupo de degustadores para avaliar a rancidez oxidativa em amostras de alimentos. Por fim, Sivertsen e Risvik (1994) combinaram PCA com GPA e ANOVA para analisar diferenças entre vinhos produzidos em regiões diferentes e avaliar o desempenho dos avaliadores.

A Análise Generalizada de Procustes ou *Generalized Procustes Analysis* (GPA) é uma técnica estatística que compara o formato de diferentes objetos. Essa comparação é feita através da translação dos dados de cada degustador para uma origem comum (neutralizando o efeito de notas individuais nos níveis da escala), de uma rotação e, por fim, de uma contração ou expansão (para ajustar diferenças de escala entre os degustadores) (SINESIO et al., 1990). Essa técnica pode ser utilizada para comparar degustadores, revelando aqueles que estão percebendo as amostras diferentes do resto do painel (BANFIELD e HARRIES, 1975), ou para comparar o desempenho de diferentes painéis avaliando produtos iguais (MCEWAN et al., 2002).

Outras técnicas utilizam conceitos de clusterização para avaliar painéis sensoriais. A clusterização de variáveis em torno de componentes latentes cria um grupo de *clusters* descrito por uma variável latente de tal forma que a variância em cada grupo seja minimizada. Ao analisar o componente latente de cada *cluster*, pode-se comparar os atributos de determinado produto (VIGNEAU et al., 2003). No entanto, essa análise só é feita em duas dimensões (atributos \times produtos), sendo necessária fazer a média entre os degustadores, perdendo essa informação. Para suprir esta carência, Wilderjans e Cariou (2016) propuseram o método CLV3W, uma extensão do CLV para dados em três dimensões (degustadores \times atributos \times produtos); tais autores usaram esta técnica para simultaneamente comparar os atributos e o desempenho de degustadores.

Outro modelo utilizado para espaços de mais de duas dimensões é a análise de fatores paralelos, *Parallel factor analysis* (PARAFAC). Trata-se de um método de decomposição multivariado originado no campo de psicometria e que pode ser visto como uma generalização da análise de componentes principais (PCA) para ordens superiores (BRO, 1997). O modelo gera parâmetros que representam a variabilidade nas dimensões de interesse (por exemplo, produtos, degustadores e atributos) e permitem sua análise. Cocchi et al. (2006) utilizaram o PARAFAC para analisar os efeitos do envelhecimento de amostras de Aceto Balsâmico Tradizionale di Modena (ABTM) sobre suas características sensoriais.

Por fim, também encontram espaço na literatura abordagens apoiadas no cálculo de um índice único para avaliar os desempenhos de um painel e de cada degustador. Um desses índices é o Alfa de Cronbach descrito por Pinto et al. (2014) e baseado no coeficiente Alfa de Cronbach (Cronbach, 1951), o qual é calculado utilizando a

variância total de cada avaliação e a covariância média entre avaliações. Já Ledauphin et al. (2006) propuseram um coeficiente calculado através do autovetor correspondente ao maior autovalor de uma matriz de similaridade obtida através da comparação par a par das avaliações normalizadas de cada degustador, sendo que este indicador varia entre +1 e -1, com +1 representando a consistência total (do degustador ou do painel) e -1 a discordância total. Por se tratar do índice abordado neste estudo, a próxima seção traz a fundamentação matemática e premissas do índice.

2.2.2. Alfa de Ledauphin

Dado um painel sensorial no qual I degustadores avaliam N produtos em M atributos, pode-se montar I matrizes do tipo \mathbf{Y}_i com cada matriz tendo uma dimensão $n \times m$. Com essa notação, o elemento Y_{im} representa a avaliação feita pelo degustador i em um produto n em relação ao atributo m .

O primeiro passo para obtenção do índice é remover os vieses das avaliações dos degustadores. Para remover o efeito da utilização de diferentes níveis da escala (tipicamente verificado em avaliações), o primeiro procedimento consiste em centrar na origem cada coluna referente a cada degustador. Isso é realizado subtraindo a média de cada coluna, gerando as matrizes centradas $\mathbf{YC}_1, \mathbf{YC}_2, \dots, \mathbf{YC}_I$. Outra possível fonte de erro entre os degustadores é uso de intervalos diferentes da escala. Para corrigir isso é necessário multiplicar cada matriz centrada por um escalar θ_i , tal que $\theta_i = \frac{1}{\sqrt{f_i}}$, onde f_i é a soma dos quadrados de todos os elementos de cada matriz centrada \mathbf{YC}_i .

Com a multiplicação de cada matriz \mathbf{YC}_i por θ_i , geram-se as I matrizes corrigidas \mathbf{Z}_i . O próximo passo é comparar cada par de matrizes \mathbf{Z}_i através do cálculo de d_{i1i2} , como na equação (1).

$$d_{i1i2} = \text{traço}(\mathbf{Z}_{i1}^T \mathbf{Z}_{i2}) \quad (1)$$

Com o cálculo de d_{i1i2} , pode-se definir a similaridade entre cada par de degustadores por s_{i1i2} , com $s_{i1i2} = \frac{1+d_{i1i2}}{2}$, podendo d_{i1i2} ser entendido como a correlação entre as matrizes \mathbf{Z}_{i1} e \mathbf{Z}_{i2} consideradas como vetores pelo rearranjo dos termos em uma só coluna. Por isso, d tem características similares ao coeficiente de correlação, variando entre -1 e +1, com s variando entre 0 e 1. Quando $s = 0$ ($d = -1$), os degustadores estão em total desacordo, de tal forma que $\mathbf{Z}_{i1} = -\mathbf{Z}_{i2}$. De maneira

semelhante, $s = 1$ ($d = +1$) ocorre quando os degustadores estão em total acordo, tal que $\mathbf{Z}_{i1} = \mathbf{Z}_{i2}$. Por fim, quando não há correlação entre \mathbf{Z}_{i1} e \mathbf{Z}_{i2} , $s = 0,5$ ($d = 0$).

Através dos valores de s pode-se montar a matriz \mathbf{S} , de dimensão $I \times I$, com cada elemento $(i1, i2)$ da matriz sendo igual ao valor de s_{i1i2} . Essa matriz é simétrica, pois $s_{i1i2} = s_{i2i1}$, e cujos termos da diagonal principal são iguais a 1 (pois $s_{i1i2} = 1$ sempre que $i1 = i2$). Deve-se então extrair de \mathbf{S} o autovetor $\beta = (\beta_1, \beta_2, \dots, \beta_I)^T$ correspondente ao maior autovalor de \mathbf{S} . Esse vetor β deve ser normalizado, de forma que a soma de todos os seus elementos β_i seja igual a 1. Os coeficientes β_i variam entre 0 e 1 e representam o grau de concordância do degustador i com o restante do painel. Assim, é possível definir a matriz \mathbf{C} , que representa uma média ponderada de notas do painel, através da equação (2).

$$\mathbf{C} = \sum_{i=1}^I \beta_i \mathbf{Z}_i \quad (2)$$

A partir do cálculo da matriz \mathbf{C} , é possível, através da equação (3), calcular um indicador α_i que avalia o desempenho de cada degustador.

$$\alpha_i = \frac{\text{traço}(\mathbf{Z}_i^T \mathbf{C})}{\sqrt{\text{traço}(\mathbf{C}^T \mathbf{C})}} \quad (3)$$

O coeficiente α_i pode ser entendido como o coeficiente de correlação entre as matrizes \mathbf{Z}_i e \mathbf{C} consideradas como vetores e, portanto, pode variar entre -1 e $+1$. Um valor negativo de α_i indica que o degustador i julga os produtos de forma substancialmente diversa dos demais degustadores. De forma contrária, quanto mais próximo α_i estiver de 1, mais o degustador está em acordo com o grupo. Através da média dos valores de α_i pode-se obter um índice de desempenho global para o painel, γ .

Quanto maior o valor de γ , maior é a concordância do painel, sendo que este índice só pode ser igual a 1 se, e somente se, todas as matrizes \mathbf{Z}_i forem iguais. Além de avaliar o desempenho do painel na sua configuração original, γ pode também ser utilizado para avaliar como se altera o desempenho do painel com alterações nos dados (retirada de degustador, de atributos ou mudanças de notas).

2.3. Procedimentos metodológicos

O objetivo deste estudo é analisar o efeito esperado do treinamento de degustadores sobre o desempenho do painel sensorial através do Índice Alfa de Ledauphin. Serão comparadas três alternativas de treinamento: treinamento de um

atributo específico para todos os degustadores (treinamento dos atributos), treinamento de um determinado degustador para todos os atributos (treinamento dos degustadores) e treinamento conjunto de um atributo específico para determinado degustador (treinamento combinado).

Como notação, considera-se um painel sensorial no qual I degustadores avaliam N produtos em M atributos, gerando I matrizes \mathbf{Y}_i de dimensão $N \times M$ nas quais as linhas representam os produtos e as colunas representam os atributos. Para o cálculo dos efeitos do treinamento é necessário gerar uma matriz de notas ponderada, \mathbf{L} , calculada pela ponderação de cada uma das I matrizes de notas de cada degustador pelos autovetores normalizados, $\beta = (\beta_1, \beta_2, \dots, \beta_I)^T$, conforme a equação (4).

$$\mathbf{L} = \sum_{i=1}^I \beta_i \mathbf{Y}_i \quad (4)$$

Percebe-se que o cálculo de \mathbf{L} é análogo ao cálculo da matriz \mathbf{C} apresentada na seção 2.2.1, com a diferença de que \mathbf{C} utiliza as matrizes corrigidas \mathbf{Z}_i , ao passo que \mathbf{L} utiliza as matrizes originais \mathbf{Y}_i . Para simular o efeito do treinamento de um degustador, parte-se da premissa que o treinamento tende a aproximar a nota do degustador da nota ponderada do painel, sendo a magnitude dessa aproximação dada pela qualidade do treinamento. Definindo it como índice de treinamento (variando entre 0 e 1) e \mathbf{W} como a matriz de avaliações após treinamento, obtêm-se \mathbf{W} através da equação (5):

$$\mathbf{W} = (1 - it) * \mathbf{Y} + it * \mathbf{L} \quad (5)$$

Pela equação 5, pode-se verificar que quanto maior o valor de it , mais a nota do degustador se aproxima da nota ponderada do painel após treinamento. Para a comparação do desempenho das diferentes alternativas de treinamento, este artigo testará os valores de it iguais a 1, 0,75, 0,5 e 0,25 (quando menor it , menor a assimilação do treinamento por parte do degustador).

Definido o procedimento para incorporação do aprendizado do degustador, parte-se para as etapas remanescentes do método, as quais apoiam-se no: (i) cálculo do Alfa de Ledauphin inicial; (ii) treinamento dos atributos; (iii) treinamento dos degustadores; (iv) treinamento combinado, e (v) comparação dos resultados. Vale ressaltar que as etapas (ii), (iii) e (iv) são realizadas de forma paralela e não sequencial; depois de simular o treinamento dos atributos na etapa (ii), retoma-se o banco inicial para a etapa (iii) e o mesmo ocorre para a etapa (iv). Tais etapas são agora detalhadas.

A primeira etapa tem como objetivo calcular o Alfa de Ledauphin, índice que será utilizado como medida de desempenho do painel e que é obtido através da média dos índices α_i de cada degustador. Além disso, nessa etapa é calculada a matriz \mathbf{L} , que objetiva simular o efeito do treinamento sobre as notas dos degustadores. Os índices α_i e β_i obtidos nessa seção são calculados através dos procedimentos descritos na seção 2.2.2. Essa etapa cumpre a função de estabelecer valores de referência para as próximas análises.

A segunda etapa visa a avaliar o efeito de treinamento dos atributos para todos os degustadores e sugerir a melhor sequência de treinamento (sequência de atributos que maximiza o desempenho do painel para uma determinada quantidade de treinamentos). O treinamento de um atributo m e a consideração das potenciais melhorias em termos de homogeneização dos pareceres emitidos pelos degustadores treinados é realizado pela substituição, para cada degustador, das notas conferidas a um atributo, conforme a equação (6), na qual it (índice de treinamento) varia entre 0 e 1 e representa o quanto o treinamento do painel foi efetivo (quando it é 1, a nota do degustador torna-se igual a nota ponderada do painel):

$$\mathbf{W}_m = (1 - it) * \mathbf{Y}_m + it * \mathbf{L}_m \quad (6)$$

Inicialmente faz-se o treinamento de cada um dos M atributos, um de cada vez, avaliando o desempenho desse treinamento pelo cálculo do Índice Alfa do painel e sua comparação com o índice inicial obtido na etapa (i). Define-se então o atributo que gerou o maior alfa como m_1 , o qual deve ser o primeiro a ser treinado nos degustadores para maximizar o desempenho do painel.

Na sequência, substitui-se permanentemente o atributo m_1 nas matrizes de cada degustador, e realiza-se o treinamento de cada um dos $M - 1$ atributos restantes. Cada treinamento tem seu desempenho avaliado e define-se m_2 como o atributo que gerou o melhor resultado. Dessa forma, m_1 e m_2 representam os dois atributos que devem ser treinados para se obter a maior consistência do painel. A próxima etapa então substitui os atributos m_1 e m_2 nas matrizes \mathbf{Y}_i , e avalia o treinamento dos $M - 2$ atributos restantes, gerando, de forma análoga aos casos anteriores, m_3 . Tal processo iterativo é realizado M vezes, gerando uma sequência de treinamento dos atributos m_1, m_2, \dots, m_M .

A terceira etapa, realizada de forma paralela à segunda etapa, tem o objetivo de avaliar o efeito de se treinar um degustador para todos os atributos e sugerir a melhor

sequência de degustadores a serem treinados. O treinamento de um degustador i consiste em substituir todas as suas avaliações por aquelas após treinamento, conforme equação (7)

$$\mathbf{W}_i = (1 - it) * \mathbf{Y}_i + it * \mathbf{L} \quad (7)$$

Treinando e avaliando o desempenho de cada um dos I degustadores, define-se o degustador que gerou o melhor desempenho como i_1 . A partir disso, substitui-se a matriz \mathbf{Y}_{i1} por \mathbf{W}_{i1} de forma definitiva e calcula-se o efeito do treinamento dos $I - 1$ degustadores restantes, obtendo i_2 . O processo então continua por I etapas, de forma análoga ao apresentado na segunda etapa, obtendo-se ao final uma sequência de degustadores a serem treinados i_1, i_2, \dots, i_I .

A etapa (iv) é similar às etapas (ii) e (iii). No entanto, ao invés de se treinar um atributo para todos os degustadores ou treinar um degustador em todos os atributos, treina-se, a cada passo, apenas um atributo para um determinado degustador. O treinamento de um atributo m para um degustador i é feito através da equação (8):

$$\mathbf{W}_{im} = (1 - it) * \mathbf{Y}_{im} + it * \mathbf{L}_m \quad (8)$$

De forma semelhante às etapas anteriores, escolhe-se então a combinação de atributo e degustador que gerou o maior Alfa, substitui-se esse valor permanentemente e testam-se as $(M \times I - 1)$ combinações restantes. Esse processo ocorre $M \times I$ vezes, até que todos os valores tenham sido substituídos.

Por fim, a etapa (v) tem como objetivo comparar os resultados das etapas (ii), (iii) e (iv), avaliando o tipo de treinamento que resulta no melhor resultado (avaliado através do aumento do Alfa de Ledauphin). Os resultados serão comparados para quatro valores de índice de treinamento: 1, 0,75, 0,5 e 0,25. Para que a comparação seja adequada, serão comparados também o número de valores substituídos nas matrizes \mathbf{X} para cada treinamento; o treinamento de um atributo para todos os degustadores substitui $I \times N$ valores, o treinamento de um degustador para todos os atributos substitui $M \times N$ valores e o treinamento de apenas um atributo para um degustador envolve N valores.

2.4. Resultados

O banco de dados analisado refere-se a um painel sensorial de amostras de café especiais processados por via-úmida. Nele, 11 degustadores avaliam 9 produtos

descritos por 10 atributos: fragrância / aroma, uniformidade, ausência de defeitos, doçura, sabor, acidez, corpo, finalização, equilíbrio e final, atribuindo notas de 0 a 10, numa escala do tipo maior-é-melhor. Os 11 degustadores são certificados pelo Coffee Quality Institute (CEI) e o painel sensorial seguiu o protocolo da Specialty Coffee Association of America (SCA, 2018).

Na sequência são apresentados os resultados obtidos com a aplicação do método descrito na seção 2.3. Essa seção segue uma divisão por subseções de acordo com as cinco etapas do método proposto.

2.4.1 Cálculo do Alfa de Ledauphin inicial

Aplicando os procedimentos descritos na etapa (i), obteve-se para, cada degustador, o seu α_i , e o autovetor associado, β_i , utilizado para cálculo da matriz ponderada \mathbf{Z} , conforme apresentado na Tabela 2.1.

Tabela 2.1: Índices Iniciais

Degustador	Alfa	Beta
1	0,748	0,097
2	0,768	0,098
3	0,529	0,088
4	0,539	0,089
5	0,496	0,087
6	0,763	0,098
7	0,674	0,094
8	0,494	0,087
9	0,538	0,089
10	0,664	0,094
11	0,282	0,078
Geral	0,590	-

Através do Alfa calculado para cada degustador, encontra-se o Alfa médio do painel, 0,590, número que será utilizado como parâmetro de referência nas etapas seguintes. Pode-se perceber que o degustador 11 tem um desempenho substancialmente baixo quando comparado aos demais integrantes do painel (Alfa = 0,282), sugerindo a falta de concordância do mesmo com o restante do painel. Por outro lado, o degustador com melhor desempenho é o degustador 2, com um índice de 0,763. Analisando os valores de Beta, pode-se perceber que eles são proporcionais ao valor de Alfa, com o

pior avaliador apresentando o menor Beta e o melhor apresentando o Beta mais elevado. Isso torna a ponderação menos sujeita a erros causados por degustadores inconsistentes.

2.4.2 Treinamento dos atributos

Seguindo os procedimentos da etapa (ii), foi realizado o treinamento dos atributos com um total de 10 etapas, fixando a cada etapa a configuração anterior e selecionando o melhor atributo a ser gerado. A Tabela 2.2 traz o Alfa obtido com o treinamento de cada atributo a cada etapa, utilizando it igual a 1. Depois que um atributo é escolhido como o melhor de determinada etapa, ele não é mais treinado, o que é representado por “-” na tabela.

Com base em tais resultados, percebe-se que o Alfa cresce a cada etapa, chegando, como esperado, ao valor 1 no final do treinamento. A ordem ótima de atributos a ser treinada é: doçura, ausência de defeitos, acidez, sabor, uniformidade, final, finalização, fragrância/aroma e equilíbrio. Os resultados da Tabela 2.2 podem ser utilizados pelo gestor do painel sensorial na obtenção de um Alfa-meta ou para treinar apenas um número definido de atributos (dada uma limitação de recursos).

Tabela 2.2: Treinamento dos atributos

Atributo	Etapa 1	Etapa 2	Etapa 3	Etapa 4	Etapa 5	Etapa 6	Etapa 7	Etapa 8	Etapa 9	Etapa 10
Fragrância/aroma	0,610	0,667	0,713	0,748	0,784	0,820	0,862	0,906	-	-
Uniformidade	0,608	0,666	0,715	0,751	0,788	-	-	-	-	-
Ausência de defeitos	0,626	0,689	-	-	-	-	-	-	-	-
Doçura	0,645	-	-	-	-	-	-	-	-	-
Sabor	0,614	0,672	0,720	0,755	-	-	-	-	-	-
Acidez	0,616	0,674	0,722	-	-	-	-	-	-	-
Corpo	0,609	0,666	0,712	0,747	0,782	0,820	0,861	0,904	0,951	-
Finalização	0,611	0,669	0,716	0,751	0,787	0,823	0,865	-	-	-
Equilíbrio	0,607	0,663	0,710	0,745	0,780	0,816	0,857	0,902	0,948	1,000
Final	0,610	0,668	0,716	0,752	0,788	0,826	-	-	-	-
Máximo	0,645	0,689	0,722	0,755	0,788	0,826	0,865	0,906	0,951	1,000

No caso de se buscar um Alfa-meta, encontra-se na tabela o primeiro valor de Alfa maior que a meta e faz-se o treinamento dos atributos dessa etapa e das etapas anteriores. Para a situação em que há um limite de atributos a serem treinados, simplesmente treina-se os atributos das etapas 1 até o limite.

2.4.3 Treinamento dos degustadores

De forma análoga a seção 2.4.2, seguiram-se os procedimentos da etapa (iii) para a avaliação do impacto de se treinar um determinado degustador para todos os atributos sobre o desempenho do painel. Esse treinamento ocorreu em 11 etapas, valor igual a

quantidade de degustadores. Os resultados são apresentados na Tabela 2.3, que segue o mesmo formato da Tabela 2.2.

Pode-se perceber, novamente, que o valor de Alfa aumenta a cada etapa, com o valor final sendo igual a 1. A ordem dos degustadores a ser treinada é: 11, 5, 8, 4, 3, 9, 7, 10, 1, 2 e 6. Vale ressaltar que essa ordem de degustadores é similar, mas não igual, à ordem decrescente de degustadores por seu índice Alfa de Ledauphin (conforme apresentado na Tabela 2.1), o que justifica o uso dessa ferramenta.

Os resultados da Tabela 2.3 podem ser utilizados pelo gestor do painel de forma semelhante aos resultados da Tabela 2.2, para a definição dos degustadores prioritários com vistas a atingir-se um Alfa-meta ou uma limitação de degustadores.

Tabela 2.3: Treinamento dos degustadores

Degustador	Etapa 1	Etapa 2	Etapa 3	Etapa 4	Etapa 5	Etapa 6	Etapa 7	Etapa 8	Etapa 9	Etapa 10	Etapa 11
1	0,615	0,685	0,737	0,783	0,826	0,865	0,902	0,935	0,966	-	-
2	0,613	0,681	0,731	0,779	0,823	0,861	0,899	0,932	0,962	0,984	-
3	0,637	0,703	0,754	0,799	0,841	-	-	-	-	-	-
4	0,636	0,704	0,756	0,802	-	-	-	-	-	-	-
5	0,640	0,710	-	-	-	-	-	-	-	-	-
6	0,614	0,681	0,730	0,776	0,821	0,861	0,898	0,931	0,962	0,983	1,000
7	0,624	0,692	0,744	0,790	0,836	0,874	0,913	-	-	-	-
8	0,639	0,708	0,758	-	-	-	-	-	-	-	-
9	0,636	0,702	0,752	0,798	0,841	0,880	-	-	-	-	-
10	0,625	0,691	0,742	0,787	0,832	0,871	0,909	0,944	-	-	-
11	0,659	-	-	-	-	-	-	-	-	-	-
Máximo	0,659	0,710	0,758	0,802	0,841	0,880	0,913	0,944	0,966	0,984	1,000

2.4.4 Treinamento combinado

Nessa etapa propõe-se o treinamento de apenas um atributo de um degustador. Esse treinamento apoia-se em 110 etapas, geradas pela multiplicação do número de degustadores (11) pela quantidade de atributos (10). A partir desse treinamento, gera-se uma tabela semelhante às Tabelas 2.2 e 2.3, com cada linha representando uma combinação entre degustador e atributo. A Tabela 2.4 ilustra os resultados gerados com a utilização de *it* igual a 1; a tabela em seu tamanho original é suprimida por limitações de espaço.

Tabela 2.4: Treinamento combinado - parcial

Degustador	Atributo	Etapa 1	Etapa 2	Etapa 3	...	Etapa 74	...	Etapa 108	Etapa 109	Etapa 110
1	Fragrância/aroma	0,592	0,606	0,615	...	0,000	...	-	-	-
2	Fragrância/aroma	0,589	0,603	0,612	...	0,942	...	-	-	-
3	Fragrância/aroma	0,592	0,606	0,615	...	0,000	...	-	-	-
4	Fragrância/aroma	0,593	0,607	0,616	...	0,000	...	-	-	-
5	Fragrância/aroma	0,599	0,613	0,622	...	0,000	...	-	-	-
6	Fragrância/aroma	0,589	0,603	0,612	...	0,941	...	0,993	-	-
...
6	Finalização	0,590	0,604	0,613	...	0,942	...	0,991	0,995	1,000
7	Finalização	0,590	0,604	0,614	...	0,943	...	-	-	-
...
11	Final	0,596	0,609	0,618	...	0,000	...	-	-	-
	Máximo	0,599	0,613	0,622	0,636	0,943	0,993	0,995	1,000	1,000

Assim como nos casos anteriores, o valor de Alfa aumenta a cada etapa e o valor após a última etapa é igual a 1. A tabela 2.4 completa pode ser utilizada pelo gestor do painel para alcançar os mesmos objetivos já explicados nas etapas anteriores.

2.4.5 Comparação dos resultados

Esta etapa visa a identificar o tipo de treinamento – atributo de forma isolada, degustador de forma isolada ou combinado – que demanda um menor número de substituições nas avaliações (representando o efeito de treinamentos). Cada etapa de treinamento por atributo gera a necessidade de 99 substituições (11×9), cada etapa do treinamento do degustador necessita de 90 substituições (10×9) e cada etapa do combinado exige 9 substituições.

No entanto, para uma comparação correta, seria necessário manter o mesmo número de substituições para cada método, o que torna a comparação do treinamento por atributo e por degustador inviável. Para solucionar o problema, decidiu-se gerar uma interpolação linear entre os valores de substituição ausentes. Além disso, a necessidade de substituição de cada método foi dividida por 9 (número de produtos comum a todos os métodos para facilitar a exibição). Fazendo essa divisão e a interpolação, gerou-se a Figura 2.1, que compara os três métodos utilizando um valor de it igual a 1.

A Figura 2.1 permite concluir que treinamentos de forma combinada (ou seja, focando em atributos específicos de determinados degustadores) conduz a valores superiores de Alfa para quase todo o espaço de substituições. Além disso, percebe-se que o valor máximo do Alfa (quando todos os degustadores e atributos foram treinados) é igual a 1. Geram-se então as Figuras 2.2, 2.3 e 2.4, que apresentam os resultados para valores de índice de aprendizado iguais a 0,75, 0,5 e 0,25, respectivamente.

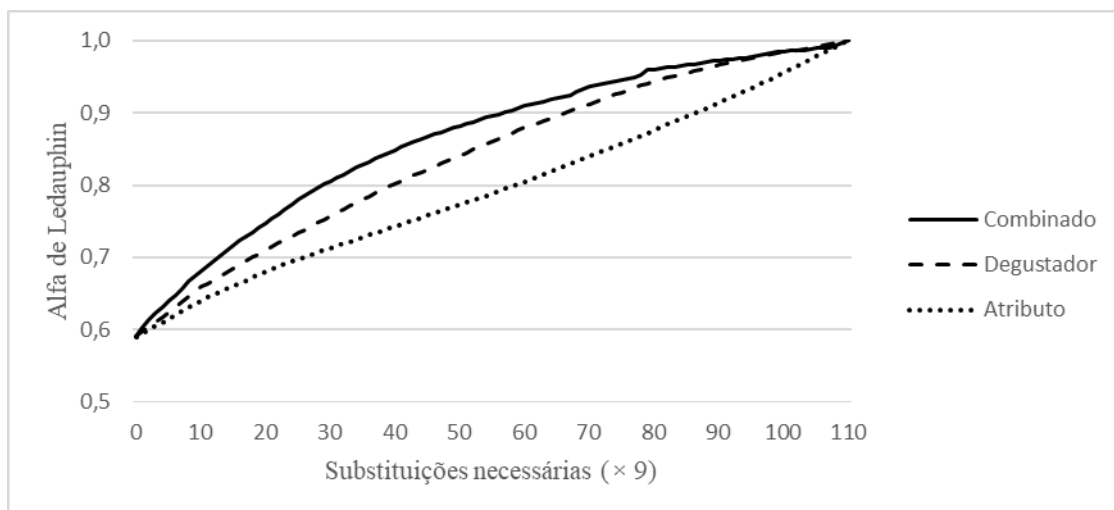


Figura 2.1: Comparação dos métodos com $it = 1$

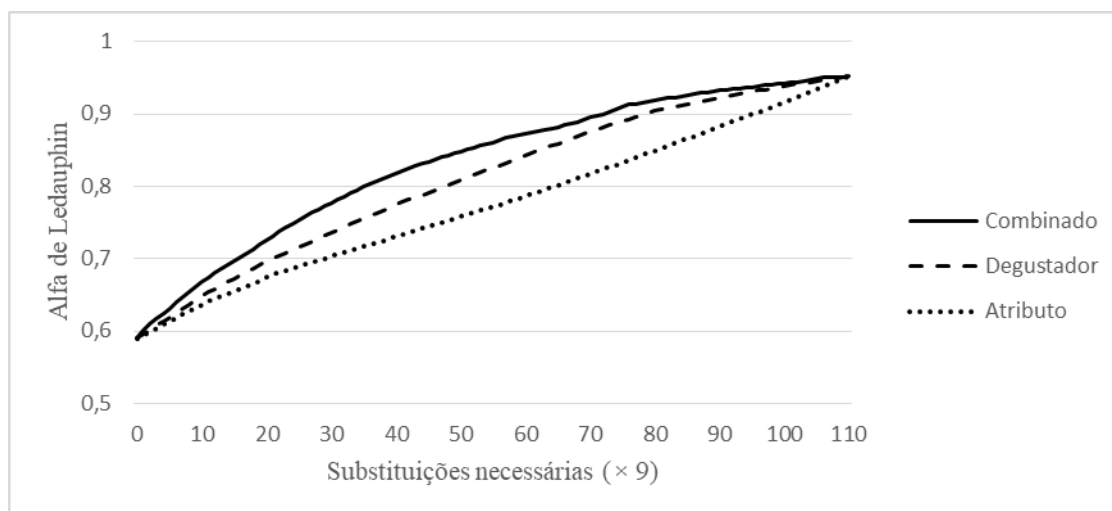


Figura 2.2: Comparação dos métodos com $it = 0,75$

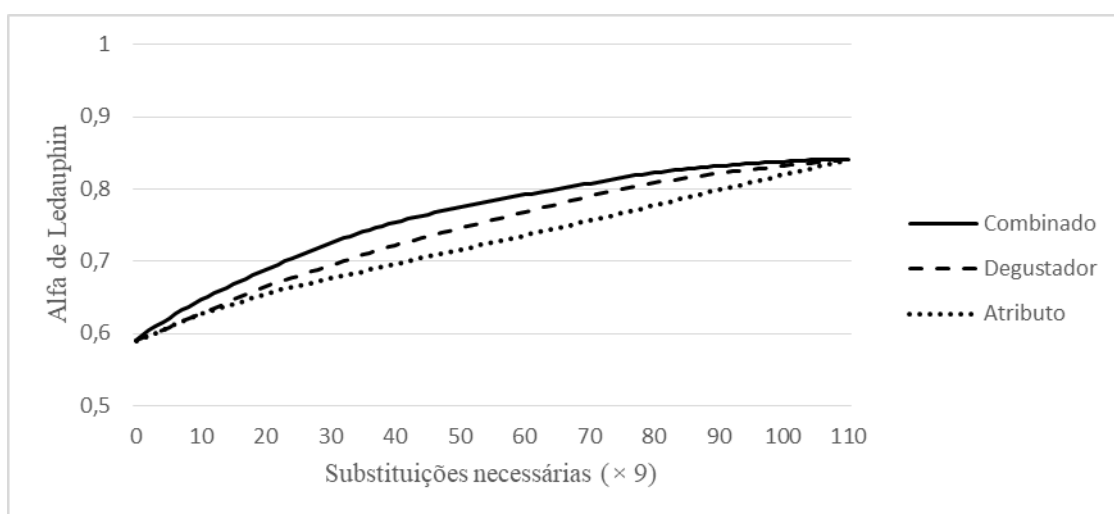


Figura 2.3: Comparação dos índices com $it = 0,5$

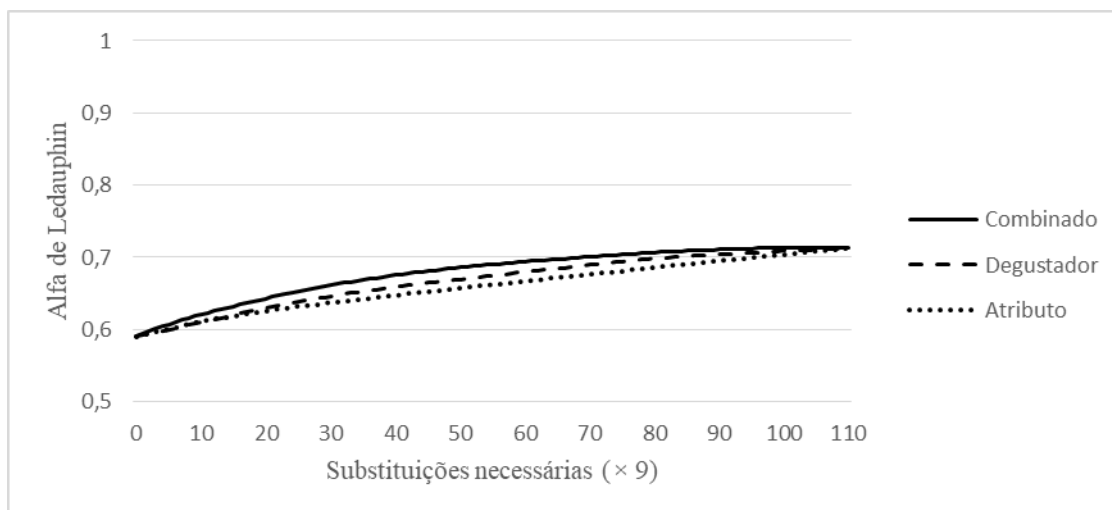


Figura 2.4: Comparação dos índices com $it = 0,25$

A primeira conclusão obtida das Figuras 2.1, 2.2, 2.3 e 2.4 é que o método combinado apresenta melhor desempenho para todos os valores de it testados. No entanto, a diferença entre o desempenho do método combinado e dos demais métodos reduz-se com a redução dos valores de it . Para visualizar esses resultados de forma numérica, montou-se a Tabela 2.5, que mostra, para cada método de treinamento, o Alfa Médio e o Alfa Máximo.

Tabela 2.5: Índices Alfa

Método	$it = 1$	$it = 0,75$	$it = 0,5$	$it = 0,25$	Média
Combinado	0,868	0,836	0,765	0,679	0,787
Degustador	0,843	0,813	0,746	0,669	0,768
Atributos	0,796	0,777	0,726	0,660	0,740
Alfa Máximo	1	0,953	0,841	0,713	-

A partir dos resultados da Tabela 2.5, pode-se concluir que o ordenamento dos métodos em função de seu desempenho mantém-se o mesmo independente do valor de it , com o treinamento combinado apresentando o melhor resultado, seguido do treinamento por degustador. Também se pode notar a redução do Alfa Máximo com a redução de it , o que é justificado pelo fato de que menores valores de it fazem com que o treinamento tenha menor efeito, e as notas dos degustadores mantenham-se mais próximas às originais.

Outra forma de se comparar o desempenho de cada método consiste em avaliar a quantidade de substituições necessárias. Para isso, parte-se do Alfa mínimo (0,590) e

vai-se até o Alfa máximo, em incrementos de 0,05, encontrando para cada valor de Alfa a quantidade de substituições que cada método necessita até atingir tal valor de Alfa (por exemplo, para um valor de it igual a 1, o método combinado necessita de 8 substituições para atingir um Alfa de 0,66, enquanto o método por degustador exige 11 e o método por atributos exige 15). Fazendo a média desses valores para cada método e valor de it gera-se a Tabela 2.6.

Tabela 2.6: Média de substituições necessárias

Método	$it = 1$	$it = 0,75$	$it = 0,5$	$it = 0,25$	Média
Combinado	37,46	37,01	35,61	32,78	35,72
Degustador	44,35	44,31	44,16	43,00	43,95
Atributos	57,02	55,44	53,14	50,70	54,07

Os resultados da Tabela 2.6 demonstram que o treinamento combinado exige, em média, 18,7% menos treinamento do que o treinamento por degustador e 33,9% a menos que o treinamento por atributos, enquanto que o treinamento por degustador necessita de 18,7% menos treinamento que o método por atributos (as diferenças entre as médias foram testadas por uma ANOVA e apresentam o valor de p menor do que 5%). Além disso, percebe-se uma redução das substituições necessárias com valores menores de it ; isso é explicado pela convergência mais rápida dos Alfas para valores menores de it , conforme pode ser visualizado nas Figuras 2.1 a 2.4.

2.5. Conclusões

O objetivo deste trabalho consistiu em comparar três métodos para se treinar degustadores de um painel sensorial, avaliando o ganho de desempenho do painel e a quantidade de treinamento necessária. Todos os três métodos, conforme esperado, melhoraram o desempenho do painel a cada iteração. No entanto, perceberam-se diferenças na quantidade de treinamento demandado por cada método para obter os mesmos ganhos.

O método proposto de treinamento combinado entre degustadores e atributos apresentou os melhores resultados para todos os valores de índice de aprendizado testados, gerando um ganho de desempenho com uma necessidade menor de treinamento. Depois desse método, o método tradicional de se treinar alguns degustadores para todos os atributos mostrou-se melhor do que a opção de treinar apenas alguns atributos para todos os degustadores.

Como o treinamento de degustadores está associado a custos financeiros e de tempo, reduzir o treinamento é uma das preocupações de gestores de painéis sensoriais. Por isso, a principal contribuição deste trabalho foi apresentar um método que permite aumentar o desempenho de um painel às custas de menor volume de treinamento.

Diversas propostas de estudos futuros são delineadas. O efeito dos treinamentos, neste estudo, foi simulado através da substituição das notas a serem treinadas pela matriz ponderada **Z** e por um índice de treinamento; sugere-se, para futuros trabalhos, utilizar as mesmas técnicas aqui apresentadas com base no efetivo treinamento dos degustadores (coletando as notas antes e depois do mesmo). Outro desdobramento consiste em aplicar e comparar as três técnicas de treinamento avaliadas (treinamento de atributos, treinamento de degustadores e treinamento combinado) para outros painéis sensoriais, tanto de café quanto de outros gêneros alimentícios. Dessa forma, pode-se dar mais robustez aos resultados deste estudo.

2.6. Referências

ANZANELLO, M.; FOGLIATTO, F.; ROSSINI, K. Data mining-based method for identifying discriminant attributes in sensory profiling. **Food Quality and Preference**, Volume 22, 2011.

BANFIELD, C; HARRIES, J. A technique for comparing judges' performance in sensory tests. **International Journal of Food Science & Technology**, Volume 10, 1975.

BI, J. Agreement and reliability assessments for performance of sensor descriptive panel. **Journal of Sensory Studies**, Volume 18, 2007.

BORGOGNONE, M.; BUSSI, J.; HOUGH, G. Principal component analysis in sensory analysis: covariance or correlation matrix? **Food Quality and Preference**, Volume 12, 2001.

BRO, R. PARAFAC. Tutorial and applications. **Chemometrics and Intelligent Laboratory Systems**, Volume 38, 1997.

COCCHI, M.; BRO, R.; DURANTE, C.; MANZINI, D.; MARCHETTI, A.; SACCANI, F.; SIGHINOLFI, S.; ULRICI, A. Analysis of sensory data of Aceto Balsamico Tradizionale di Modena (ABTM) of different ageing by application of PARAFAC models. **Food Quality and Preference**, Volume 17, 2006.

CRONBACH, L.; Coefficient alpha and the internal structure of tests. **Psychometrika**, Volume 16, 1951.

DIJKSTERHUIS, G. Assessing panel consonance. **Food Quality and Preference**, Volume 6, 1995.

LATREILLE, J.; MAUGER, E.; AMBROISINE, L.; TENENHAUS, M.; VINCENT, M.; NAVARRO, S.; GUINOT, C. Measurement of the reliability of sensory panel performances. **Food Quality and Preference**, Volume 17, 2006.

LEDAUPHIN, S.; HANAFI, M.; QANNARI, E. Assessment of the agreement among the subjects in fixed vocabulary profiling. **Food Quality and Preference**, Volume 17, 2006.

MCEWAN, J.; HUNTER, E.; VAN GEMERT, L.; LEA, P. Proficiency testing for sensory profile panels: measuring panel performance. **Food Quality and Preference**, Volume 13, 2002.

PINTO, F.; FOGLIATTO, F.; QANNARI, E. A method for panelists' consistency assessment in sensory evaluations based on the Cronbach's alpha coefficient. **Food Quality and Preference**, Volume 32, 2014.

ROSSI, F. Assessing sensory panelist performance using repeatability and reproducibility measures. **Food Quality and Preference**, Volume 12, 2001.

SCA, 2018. **Specialty Coffee Association Protocols**. 2018. Disponível em [≤https://sca.coffee/research/protocols-best-practices>](https://sca.coffee/research/protocols-best-practices). Acesso em 4 set. 2018.

SINESIO, F.; RISVIK, E.; RODBOTTEN, M. Evaluation of panelist performance in descriptive profiling of rancid sausages: a multivariate study. **Journal of Sensory Studies**, Volume 5, 1990.

SIVERTSEN, H.; RISVIK, E. A Study of Sample and Assessor variation: A multivariate study of wine profiles. **Journal of Sensory Studies**, Volume 9, 1994.

TALSMA, P. Assessing sensory panel performance using generalizability theory. **Food Quality and Preference**, Volume 47, 2016

VIGNEAU, E.; QANNARI, E. Clustering of variables around Latent Components. **Communications in Statistics- Simulation and Computation**, Volume 32, 2003.

WILDERJANS, T.; CARIOU, V. CLV3W: A clustering around latent variables approach to detect panel disagreement in three-way conventional sensory profiling data. **Food Quality and Preference**, Volume 47, 2016

3. Segundo artigo: Previsão de notas de atributos em um painel sensorial de café apoiada em PLS e técnicas de seleção

Luiz Antônio Nonenmacher Júnior (ljuniornone@gmail.com)

Michel José Anzanello (anzanello@producao.ufrgs.br)

Resumo

Uma das típicas funções dos painéis sensoriais na indústria de alimentos é classificar produtos em categorias de qualidade, que impactam diretamente no seu preço de venda e reconhecimento. Por isso, torna-se interessante prever a nota que um produto irá receber com base em variáveis que o descrevam. O objetivo deste estudo é utilizar a regressão PLS para prever a nota dada pelos degustadores a cada atributo de amostras de café especial, selecionar variáveis (físico-químicas e de processo) e degustadores que minimizem o erro de predição e reduzam o custo do experimento. Para isso, inicialmente faz-se a seleção de degustadores utilizando o Índice Alfa de Ledauphin e depois a seleção de variáveis com um índice de importância de variáveis baseado em PLS. O principal resultado foi a queda no número de degustadores (de 10 para 5) e no número de variáveis (de 11 para 3: altitude, pH e tipo de fermentação) com uma redução de 3,5% no erro de predição.

Palavras-chave: Painéis Sensoriais, Seleção de Variáveis, PLS, Índice Alfa de Ledauphin.

Abstract

Sensory panels are commonly utilized in the food industry for purposes of quality control, classification and research. One of the most widely use of those panels is the classification of the food samples into quality categories that directly influence the product price and its recognition. Because of that, the prediction of grades given to a product base on its variables is of interest. The objective of this study is the use of the PLS regression to predict the grades given by graders to each attribute of each special coffee sample, choosing variables and graders that minimize the prediction error and reduce the cost of the experiment. To this end, firstly the selection of variables is made utilizing the Alpha Index of Ledauphin and, thereafter, the selection of graders is made using a variable importance index based on PLS. The most important result of this study

is the reduction of evaluators (from 10 to 5: altitude, pH and type of fermentation) and variables (from 11 to 3) with the reduction of 3.5% in the prediction error.

Keywords: Sensory Panels, Variables Selection, PLS, Alpha Index of Ledauphin.

3.1. Introdução

Técnicas de análise sensorial utilizam um painel de degustadores treinados para obter um grupo de características sensoriais de um produto (WILDERJANS e CARIOU, 2016), sendo tipicamente utilizadas na indústria de alimentos para fins de controle de qualidade, classificação e pesquisa (BANFIELD e HARRIES, 1975). Um painel é consistido por um grupo de degustadores treinados que julgam amostras de produtos (com a presença ou não de repetições) com base em um conjunto de atributos, sendo os resultados tipicamente expressos numa escala numérica (ANZANELLO et al., 2011).

Um dos objetivos mais comuns em painéis sensoriais é a classificação das amostras de alimentos em categorias de qualidade, de acordo com a média das notas em cada atributo. Essa classificação tem fundamental importância para os produtores do alimento, visto que impacta diretamente no preço de venda e no reconhecimento do seu produto. Por isso, torna-se interessante prever a nota de qualidade de um produto com base em variáveis químicas ou de processo. Além disso, ao utilizar um painel de degustadores, torna-se importante definir apropriadamente a quantidade de amostras, de atributos avaliados e o número de degustadores. Utilizar uma quantidade elevada de amostras, atributos e degustadores pode gerar mais robustez e confiabilidade aos resultados; no entanto, aumenta o tempo requerido para execução do experimento e, conseqüentemente, o custo do mesmo.

O objetivo deste artigo é aplicar a regressão PLS em um banco de dados de amostras de café, com o objetivo de prever as notas conferidas a diversos atributos de relevância na qualificação do café (como acidez e doçura, dentre outros). As variáveis independentes da regressão são variáveis de processo (classe de altitude e tipo de fermentação) e físico-químicas (extrato aquoso, ácido gálico, % de acidez e pH) das amostras. Após a primeira rodada da regressão, será realizada a seleção dos degustadores mais consistentes através do Índice Alfa de Ledauphin, bem como das

variáveis independentes mais relevantes. Tal esforço de seleção objetiva aumentar a precisão de previsão e reduzir o custo de futuros experimentos.

O artigo está organizado da seguinte forma: após essa introdução, a seção 3.2 faz uma breve revisão sobre modelos de regressão, seleção de variáveis e avaliação de qualidade de degustadores, com destaque à regressão PLS e ao Índice Alfa de Ledauphin. Na sequência, a seção 3.3 descreve os procedimentos utilizados para atingimento dos objetivos, com os resultados da aplicação sendo descritos na seção 3.4. Por fim, a seção 3.5 arremata o trabalho, resumindo as principais conclusões do mesmo e sugerindo possibilidades para futuros trabalhos.

3.2. Referencial teórico

Esta seção está dividida em quatro subseções. A primeira subseção trata de técnicas de regressão, com destaque para a regressão PLS. A segunda subseção discute métricas para avaliar o erro de predição. Por fim, as subseções 3.2.3 e 3.2.4 tratam, respectivamente, de seleção de variáveis e de degustadores, com destaque para o índice de importância de variáveis utilizado nesse artigo e o Índice Alfa de Ledauphin.

3.2.1 Técnicas de regressão

Os resultados de um painel sensorial que avalia N amostras podem ser representados por duas matrizes: uma matriz de variáveis independentes \mathbf{X} , composta por J colunas (variáveis) e N linhas e uma matriz de variáveis de resposta \mathbf{Y} , composta por M colunas (atributos) e N linhas. Cada amostra pode então ser representada por um elemento x_{nj} e um elemento y_{nm} , nos quais n representa a amostra, j a variável e m o atributo. O objetivo de técnicas de regressão é obter uma estimativa da matriz \mathbf{Y} a partir dos valores de \mathbf{X} .

Uma das técnicas de regressão mais utilizadas é a regressão linear, que define um coeficiente para cada variável x_j a fim de minimizar o erro quadrático entre cada ponto y e seu valor previsto (POMBEIRO et al., 2017). A regressão linear pode ser simples, quando utiliza apenas uma variável x_j para prever \mathbf{Y} , ou múltipla, ao utilizar mais de uma variável x_j . O método, no entanto, tem como resposta apenas uma variável de resposta, sendo necessário gerar diversos modelos caso se pretenda prever mais de uma variável de saída (MONTGOMERY et al., 2012).

A principal limitação da regressão linear é lidar com banco de dados onde há muitas variáveis ou elevado grau de colinearidade entre as mesmas (BOLDINA e BENINGER, 2016). Para essas situações, utilizam-se outras técnicas que reduzem a dimensão dos dados de entrada minimizando estes problemas. Duas das técnicas mais utilizadas nesse sentido são a Regressão por Componentes Principais (*Principal Component Regression* - PCR) e a Regressão por Mínimos Quadrados Parciais (*Partial Least Squares Regression* - PLSR).

A Regressão por Componentes Principais combina a Análise de Componentes Principais (PCA) com a regressão linear. Primeiramente, através do PCA, as variáveis originais são transformadas em um conjunto reduzido de componentes principais (combinações lineares das variáveis originais) com o objetivo de explicar o máximo possível da variabilidade dos dados originais (BORGOGNONE et al., 2001). Após essa transformação, gera-se um modelo de regressão linear utilizando como variáveis independentes os dados de entrada transformados ao invés dos dados originais do problema. A principal limitação do PCR consiste em fazer a transformação dos dados utilizando apenas as variáveis independentes, sem levar em consideração a relação entre elas e as variáveis de resposta, o que diminuiu sua capacidade preditiva (GODOY et al., 2014).

Para suprir tal deficiência, pode-se utilizar a Regressão por Mínimos Quadrados Parciais, técnica que realiza uma transformação similar ao PCR, mas buscando maximizar a covariância entre as combinações lineares (variáveis latentes) das matrizes \mathbf{X} e \mathbf{Y} (ABDI, 2010). As variáveis latentes t , chamadas de escores de \mathbf{X} , são formadas a partir de combinações lineares das variáveis originais da matriz \mathbf{X} associadas a pesos w , conforme equação (1). De forma análoga, são definidas as variáveis latentes u a partir da matriz \mathbf{Y} associadas a pesos c , conforme equação (2) (WOLD et al., 2001).

$$t_{na} = w_{1a}x_{n1} + w_{2a}x_{n2} + \dots + w_{Ja}x_{nJ} = \mathbf{W}'\mathbf{X} \quad (1)$$

$$u_{na} = c_{1a}y_{n1} + c_{2a}y_{n2} + \dots + c_{Ma}y_{nM} = \mathbf{C}'\mathbf{Y} \quad (2)$$

A partir desses escores definem-se os coeficientes de regressão PLS, b_{mj} , conforme a equação (3).

$$b_{mj} = \sum_a c_{ma} \sum_j w_{ja} = \mathbf{C}'\mathbf{W} \quad (3)$$

Através desses coeficientes b da matriz \mathbf{X} , estima-se o valor de um elemento y_{nm} através da equação (4), na qual e_{nm} representa o resíduo entre o valor real e o estimado.

$$y_{nm} = \sum_j b_{mj} x_{nj} + e_{nm} = \mathbf{BX} + \mathbf{E} \quad (4)$$

O algoritmo utilizado neste artigo para determinação dos parâmetros do PLS, NIPALS (*Non-linear Iterative Partial Least Squares*), é detalhado em Geladi e Kowalski (1986). O algoritmo opera de forma iterativa: os parâmetros são inicialmente definidos para a primeira variável latente, em seguida para a segunda e assim sucessivamente. A primeira variável é a que mais explica a covariância entre \mathbf{X} e \mathbf{Y} ; cada variável subsequente explica menores percentuais dessa covariância (HÖSKULDSSON, 1988).

A quantidade de variáveis latentes é dada pelo número de componentes retidos a (valor que varia entre 1 e A), sendo o número máximo de componentes (A) igual ao menor valor entre a quantidade de variáveis dependentes e independentes (colunas de \mathbf{X} e \mathbf{Y}). Uma das formas de se definir a quantidade de componentes a serem retidos é a análise gráfica, comparando-se a quantidade de componentes com um critério de avaliação do erro de predição. Uma abordagem consiste então em definir o número de componentes que minimize o erro; no entanto, alguns autores defendem que é preferível escolher um número de componentes menor se o incremento no erro foi pequeno (GELADI e KOWALSKI, 1986; HASTIE et al., 2009).

No setor alimentício, diversos trabalhos já foram realizados com o uso do PLS. Lengard e Kermit (2006) utilizaram a técnica em amostras de tomate, determinando quais características do produto tinham maior efeito na percepção do consumidor e, simultaneamente, avaliando o desempenho de cada degustador. O PLS também foi utilizado por Lombardo (2011) para prever as notas dadas por um painel de degustadores a amostras de vinho a partir das variáveis químicas das amostras.

3.2.2 Avaliação de erro de predição

Quando se utilizam modelos de regressão, é fundamental estimar o erro de predição, tanto para prever a acurácia futura do modelo quanto para avaliar e comparar alternativas de modelos ou modificações do mesmo, como quantidade de variáveis a serem utilizadas ou parâmetros a serem definidos (KOHAVI, 1995). Para se estimar o erro, é necessário definir uma métrica para medição do mesmo, como o MAE (*Mean*

absolute error), o MAPE (*Mean absolute percentage error*) ou o MSE (*Mean squared error*). O MSE, utilizado neste artigo, calcula o erro através da média das diferenças ao quadrado entre o valor previsto (\hat{y}_n) e o real (y_n), conforme equação (5).

$$\text{MSE} = \frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2 \quad (5)$$

Após escolher a métrica, é necessário definir a técnica utilizada para mensurar o erro. Uma das técnicas mais simples é a ressubstituição, que consiste em gerar o modelo de regressão utilizando todos os dados disponíveis e avaliar o erro sobre esses mesmos dados. No entanto, por utilizar os mesmos dados para gerar o modelo e avaliar o erro, essa técnica subestima o verdadeiro erro. Uma técnica para amenizar este problema é o chamado método *holdout*, que consiste em particionar a amostra em dois subconjuntos mutuamente exclusivos, chamados de treino e teste. O conjunto de treino é utilizado para gerar o modelo e o desempenho é mensurado no banco de teste (KOHAVI, 1995).

No entanto, mesmo sendo melhor que a ressubstituição, existem técnicas que estimam o erro de forma mais precisa que o *holdout*. Uma dessas técnicas é chamada de Validação Cruzada *k-fold* e consiste em dividir o banco de dados em k subgrupos de tamanho similar. Desses subgrupos, $k - 1$ são utilizados para gerar o modelo, cujo desempenho é testado no subgrupo restante. O mesmo procedimento é realizado com os k subgrupos, e o erro final é a média dos erros obtidos. O custo computacional para essa técnica cresce com o valor de k , sendo uma das escolhas mais usuais o valor de $k = 10$, por obter uma boa relação custo benefício entre precisão e custo computacional (KIM, 2009).

Uma modificação da Validação Cruzada que melhora sua precisão é a Validação Cruzada Repetida, que repete o processo de particionamento e estimação de erro q vezes, sendo o erro a média das repetições (JIANG e WANG, 2017). Outra técnica, introduzida e descrita por Efron e Tibshirani (1998), é chamada de *Bootstrap*. Ela consiste em formar uma partição no banco de dados de o amostras, sendo estas escolhidas de forma aleatória com repetição, sendo essa partição utilizada para gerar o modelo. O processo ocorre r vezes, sendo ao final coletado o erro médio de cada partição. No entanto, ao utilizar q e r tal que o custo computacional dos dois métodos seja similar, Kim (2009) determinou que o método da Validação Cruzada Repetida tende a gerar melhores resultados que o *bootstrap*.

3.2.3 Seleção de variáveis

Em muitos casos, para melhorar a capacidade preditiva de um modelo de regressão, utilizam-se técnicas de seleção de variáveis para reduzir a quantidade de variáveis independentes, removendo variáveis ruidosas ou irrelevantes. Além disso, reduzir as variáveis de um modelo torna-o menos complexo, facilitando a interpretação e entendimento do sistema que está sendo estudado (BIANCOLILLO et al., 2016). Existem diversas técnicas que podem ser utilizadas para a seleção de variáveis, conforme apresentado por Mehmood et al. (2012). Exemplos de aplicação da seleção de variáveis com PLS podem ser encontrados no trabalho de Rossini et al. (2012) ou de Wang et al. (2015).

Dentre essas técnicas, uma família de técnicas de seleção de variáveis consiste em calcular um índice de importância para cada variável, retirando, sucessivamente aquelas variáveis tidas como de menor importância pelo índice. Tal índice pode ser calculado através da mensuração do impacto que a retirada da variável tem sobre a capacidade preditiva do modelo. Outra possibilidade é utilizar os parâmetros do modelo de predição para a geração dos índices. Nesse artigo utiliza-se um índice, do tipo maior-é-melhor, baseado nos parâmetros do PLS, conforme proposto por Anzanello et al. (2009). Para uma determinada variável j , o índice v_j é obtido através da equação (6), na qual P_{Ya} representa a fração de variação de \mathbf{Y} explicada pelo componente a .

$$v_j = \frac{\sum_{a=1}^A |w_{ja}| P_{Ya}^2}{\max_{j \in J} \sum_{a=1}^A w_{ja}^2} \quad (6)$$

3.2.4 Seleção de degustadores

A capacidade preditiva de um modelo de regressão para painéis sensoriais também está diretamente ligada à qualidade dos degustadores, sendo desejado que os mesmos sejam capazes de repetir consistentemente suas avaliações ao longo do tempo e avaliar produtos de maneira semelhante aos demais membros do painel (LATREILLE et al., 2006). De forma análoga à seleção de variáveis, pode-se fazer a seleção de degustadores, removendo os mais inconsistentes, sendo necessária a utilização de um índice para determinação da qualidade de cada degustador.

Neste trabalho será utilizado o Índice Alfa proposto por Ledauphin et al. (2006), o qual pode ser utilizado para avaliar a consistência de cada degustador com o painel e

do painel como um todo. Tal índice varia entre +1 (consistência total) e -1 (discordância total). Sendo I a quantidade de degustadores de um painel, o primeiro passo para obtenção do índice consiste em dividir a matriz \mathbf{Y} em I matrizes do tipo \mathbf{Y}_i , de dimensão $N \times M$, representando as avaliações realizadas pelo degustador i .

O próximo passo é remover os vieses das avaliações dos degustador. Para isso, primeiramente deve-se centrar na origem cada coluna referente a cada degustador, subtraindo a média de cada coluna e gerando as matrizes centradas $\mathbf{Y}\mathbf{C}_1, \mathbf{Y}\mathbf{C}_2, \dots, \mathbf{Y}\mathbf{C}_I$. Na sequência, é necessário multiplicar cada matriz centrada por um escalar θ_i , tal que $\theta_i = \frac{1}{\sqrt{f_i}}$, onde f_i é a soma dos quadrados de todos os elementos de cada matriz centrada $\mathbf{Y}\mathbf{C}_i$, para corrigir o uso de intervalos diferentes da escala entre cada degustador.

Com a multiplicação de cada matriz $\mathbf{Y}\mathbf{C}_i$ por θ_i , geram-se as I matrizes corrigidas \mathbf{Z}_i , que serão utilizadas nos próximos passos dos cálculos. O próximo passo é comparar cada par de matrizes \mathbf{Z}_i através do cálculo de d_{i1i2} através da equação (7):

$$d_{i1i2} = \text{traço}(\mathbf{Z}_{i1}^T \mathbf{Z}_{i2}) \quad (7)$$

Com o cálculo de d_{i1i2} , pode-se definir a similaridade entre cada par de degustadores por s_{i1i2} , com $s_{i1i2} = \frac{1+d_{i1i2}}{2}$, podendo d_{i1i2} ser entendido como a correlação entre as matrizes \mathbf{Z}_{i1} e \mathbf{Z}_{i2} consideradas como vetores pelo rearranjo dos termos em uma só coluna. Por isso, d tem características similares ao coeficiente de correlação, variando entre -1 e +1, com s variando entre 0 e 1. Quando $s = 0$ ($d = -1$), os degustadores estão em total desacordo, de tal forma que $\mathbf{Z}_{i1} = -\mathbf{Z}_{i2}$. De maneira semelhante, $s = 1$ ($d = +1$) ocorre quando os degustadores estão em total acordo, tal que $\mathbf{Z}_{i1} = \mathbf{Z}_{i2}$. Por fim, quando não há correlação entre \mathbf{Z}_{i1} e \mathbf{Z}_{i2} , $s = 0,5$ ($d = 0$).

Através dos valores de s pode-se montar a matriz \mathbf{S} , de dimensão $I \times I$, com cada elemento $(i1, i2)$ da matriz sendo igual ao valor de s_{i1i2} . Essa matriz é simétrica, pois $s_{i1i2} = s_{i2i1}$, e cujos termos da diagonal principal são iguais a 1 (pois $s_{i1i2} = 1$ sempre que $i1 = i2$). Deve-se então extrair de \mathbf{S} o autovetor $\beta = (\beta_1, \beta_2, \dots, \beta_I)^T$ correspondente ao maior autovalor de \mathbf{S} . Esse vetor β deve ser normalizado, de forma que a soma de todos os seus elementos β_i seja igual a 1. Os coeficientes β_i variam entre 0 e 1 e representam o grau de concordância do degustador i com o restante do painel. Assim, é possível definir a matriz \mathbf{C} , que representa uma média ponderada de notas do painel, através da equação (8).

$$\mathbf{C} = \sum_{i=1}^I \beta_i \mathbf{Z}_i \quad (8)$$

A partir do cálculo da matriz \mathbf{C} , é possível, através da equação (9), calcular um indicador α_i que avalia o desempenho de cada degustador.

$$\alpha_i = \frac{\text{traço}(\mathbf{Z}_i^T \mathbf{C})}{\sqrt{\text{traço}(\mathbf{C}^T \mathbf{C})}} \quad (9)$$

O coeficiente α_p pode ser entendido como o coeficiente de correlação entre as matrizes \mathbf{Z}_i e \mathbf{C} consideradas como vetores e, portanto, pode variar entre -1 e $+1$. Um valor negativo de α_i indica que o degustador i julga os produtos de forma substancialmente diversa dos demais degustadores. De forma contrária, quanto maior a proximidade de α_i a 1 , mais o degustador está em acordo com o grupo. Através da média dos valores de α_i pode-se obter um índice de desempenho global para o painel.

3.3. Procedimentos metodológicos

O objetivo deste trabalho é utilizar a regressão PLS para fazer a previsão das notas dadas pelos degustadores a cada produto em cada um dos atributos, buscando selecionar degustadores e variáveis que minimizem o erro de predição. Além de minimizar o erro de predição, fazer a seleção de variáveis e degustadores tem também o objetivo de reduzir o custo do experimento. O procedimento proposto é dividido em três etapas: (i) Definição do número de componentes da regressão PLS; (ii) Seleção de degustadores e (iii) Seleção de variáveis.

Nas três etapas é necessária a avaliação do erro de predição. Esse erro será medido através do MSE (*Mean squared error*), que será obtido através de uma Validação Cruzada Repetida. Essa Validação Cruzada é executada com 10 subgrupos e 100 repetições, parâmetros escolhidos pela convergência do erro. A primeira etapa consiste em executar a regressão PLS variando o número de componentes retidos e avaliando o erro de predição, selecionando a quantidade de componentes com o menor erro.

A partir do número de componentes definidos na etapa (i), a segunda etapa fará uma seleção de degustadores para aumentar a precisão do modelo. Essa seleção inicia com o cálculo do Índice Alfa de Ledaphin para cada degustador, conforme apresentado na seção 3.2.4. Os degustadores serão então ordenados de forma crescente; a cada

iteração, remove-se o degustador (e suas notas) com menor índice, gera-se um novo modelo PLS com base nas notas dos degustadores remanescentes e avalia-se o erro do modelo, até que restem apenas dois degustadores. Será então gerado um gráfico comparativo e definidos os degustadores a serem mantidos na análise.

Por fim, após a definição do número de componentes e a seleção de degustadores, é realizada a seleção das variáveis que descrevem as amostras (variáveis físico-químicas e de processo). Para isso, calcula-se, para cada variável j , o índice de importância v_j , conforme apresentado na seção 3.2.3. Depois, ordenam-se as variáveis de acordo com esse índice e aplica-se sucessivamente a regressão PLS, retirando a cada rodada a variável com menor valor de índice. Após restar apenas uma variável, monta-se um gráfico que relaciona o erro de predição à quantidade de variáveis retidas.

3.4. Resultados

O banco de dados analisado refere-se a um painel sensorial de amostras de café especiais, processados por via úmida. Nele, 11 degustadores avaliam 9 produtos em 10 atributos: fragrância / aroma, uniformidade, ausência de defeitos, doçura, sabor, acidez, corpo, finalização, equilíbrio e final, atribuindo notas de 0 a 10, numa escala do tipo maior-é-melhor. O painel sensorial foi aplicado por degustadores certificados pelo CQI (*Coffee Quality Institute*) e seguiram o protocolo da *Specialty Coffee Association of America* (SCA, 2018).

Cada um dos 9 produtos é avaliado por 4 amostras, sendo cada amostra descrita por 10 variáveis, 2 variáveis de processo, 4 variáveis físico-químicas e 4 variáveis de coloração da torra. As variáveis de processo são variáveis categóricas e, para a aplicação da regressão PLS, foram substituídas por 4 variáveis binárias (2 para a altitude e 2 para o tipo de fermentação), que indicam se a amostra pertence ou não àquela categoria. O banco de dados utilizado então é composto de 396 linhas (9 produtos \times 4 amostras \times 11 degustadores), sendo dividido na matriz \mathbf{X} , de dimensão 396×10 , e na matriz \mathbf{Y} , de dimensão 396×10 .

As variáveis de processo são a classe de altitude e o tipo de fermentação. A classe de altitude é composta por três faixas: 700 a 750 metros, 750 a 800 metros e 850 a 900 metros; já o tipo de fermentação pode ser com água, com levedura ou sem fermentação. Por sua vez, as variáveis físico-químicas dizem respeito ao extrato aquoso

(% m.s), ácido gálico (g/100g de amostra), % acidez (v/p) e pH. Por fim, os aspectos de coloração incluem L^* (luminosidade), a^* (componente vermelho-verde), b^* (componente amarelo-azul) e C^* (pureza da cor).

Na sequência são apresentados os resultados obtidos com a aplicação do método descrito na seção 3.3. Essa seção segue uma divisão por subseções de acordo com as cinco etapas do método proposto.

3.4.1 Definição do número de componentes

Aplicando os procedimentos descritos na etapa (i), gerou-se um gráfico comparando a quantidade de componentes retidos com o erro de predição (Figura 3.1). A partir dos resultados retêm-se apenas um componente, obtendo-se o menor erro (MSE = 0,479) com a menor quantidade de componentes.

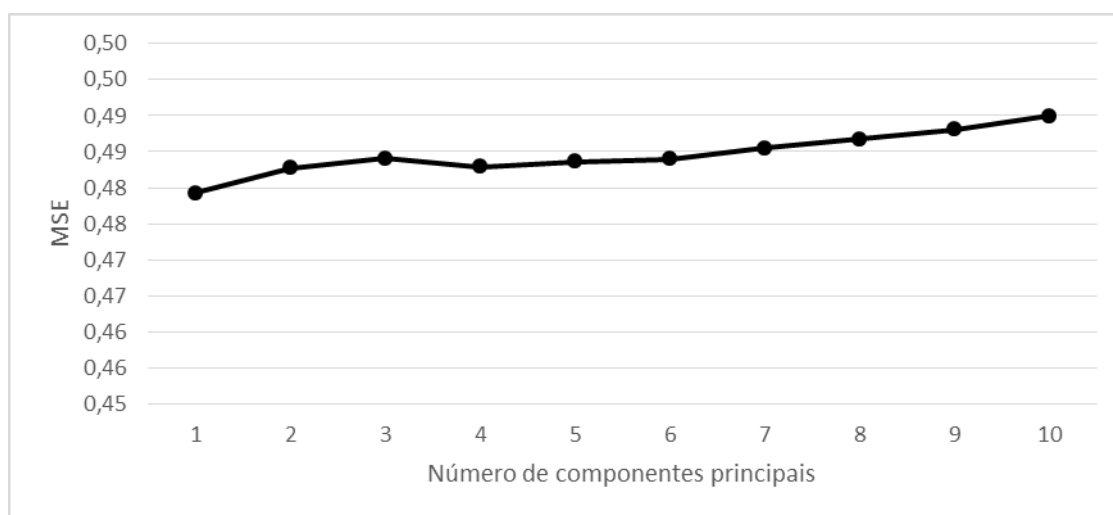


Figura 3.1: Componentes retidos

3.4.2 Seleção de degustadores

O primeiro passo dessa etapa é calcular o Alfa de Ledauphin para cada degustador e ordená-los de forma crescente. Conforme a Tabela 3.1, a ordem de degustadores a ser retirada é: 11, 8, 5, 3, 9, 4, 10, 7 e 1 (os dois últimos degustadores serão mantidos). Fazendo a retirada nesta ordem e avaliando o MSE de cada etapa, obtêm-se o gráfico mostrado na Figura 3.2.

Analisando a figura, conclui-se que o número recomendado de degustadores é cinco (degustadores 2, 6, 1, 7 e 10), situação na qual o MSE é reduzido de 0,479 para

0,466. Embora a redução do erro não seja elevada, a redução no número de degustadores reduz o custo do experimento sem prejudicar a eficácia da predição.

Tabela 3.1: Alfa de Ledauphin

Degustador	Alfa
11	0,282
8	0,494
5	0,496
3	0,529
9	0,538
4	0,539
10	0,664
7	0,674
1	0,748
6	0,763
2	0,768
Média	0,590

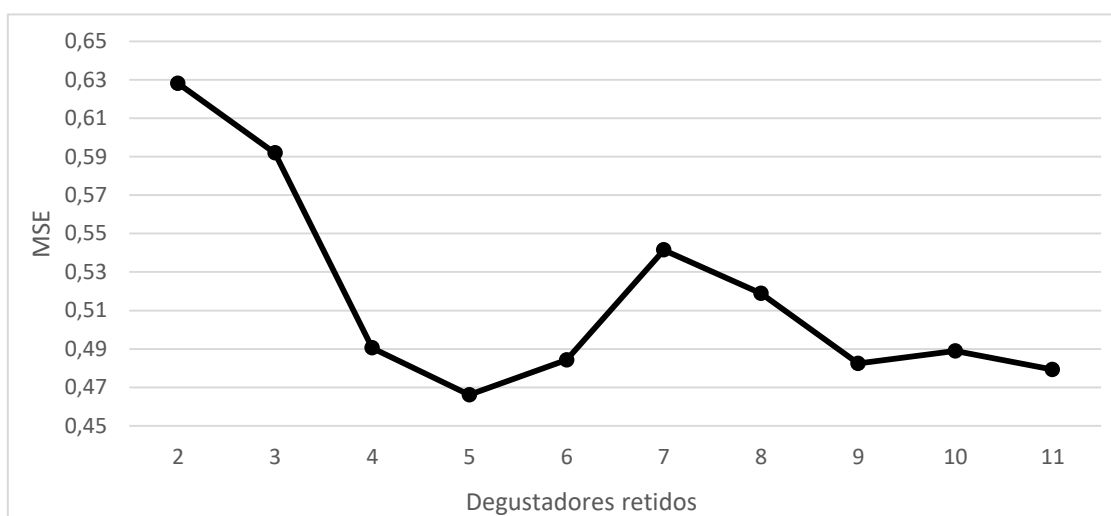


Figura 3.2: Seleção de degustadores

3.4.3 Seleção de variáveis

A partir dos cinco degustadores selecionados na etapa anterior, o próximo passo consiste em calcular o índice de seleção de variáveis para cada variável, conforme apresentado na Tabela 3.2.

Tabela 3.2: Índice de seleção de variáveis

Variável	Tipo	Índice
b*	Coloração	7,24E-06
c*	Coloração	0,0002
L*	Coloração	0,0007
a*	Coloração	0,0278
Extrato aquoso	Físico-química	0,0387
Acidez	Físico-química	0,0470
Ácido gálico	Físico-química	0,0492
Fermentação	Processo	0,0582
PH	Físico-química	0,3885
Classe altitude	Processo	1,0000

Pelo cálculo dos índices, pode-se perceber que as variáveis de processo são as que mais diferenciam as amostras, seguidas pelas físico-químicas (em especial o pH); as variáveis referentes à coloração sendo as menos importantes. Fazendo então a retirada de variáveis seguindo a ordem do índice, obtêm-se o gráfico presente na Figura 3.3.

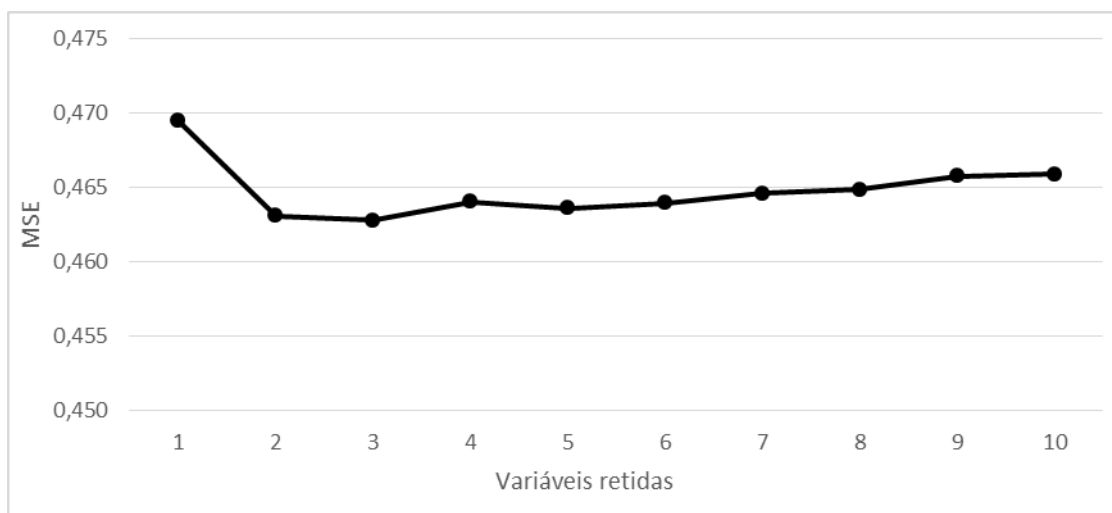


Figura 3.3: Seleção de variáveis

A Figura 3.3 sugere que o menor valor de MSE é obtido ao reter-se 3 das 10 variáveis originais (Classe de altitude, pH e Fermentação), com a redução do MSE de 0,466 para 0,463. Essa pequena redução pode ter sido causada pela incapacidade do índice de selecionar corretamente as variáveis ou pela robustez da regressão PLS frente ao número de variáveis.

Para avaliar se a causa foi o índice de seleção de variáveis, fez-se o mesmo procedimento utilizando uma técnica de enumeração total, que consiste em, testar todas

as combinações de variáveis e manter aquelas que geram o menor erro. A comparação deste método com o índice encontra-se na Figura 3.4.

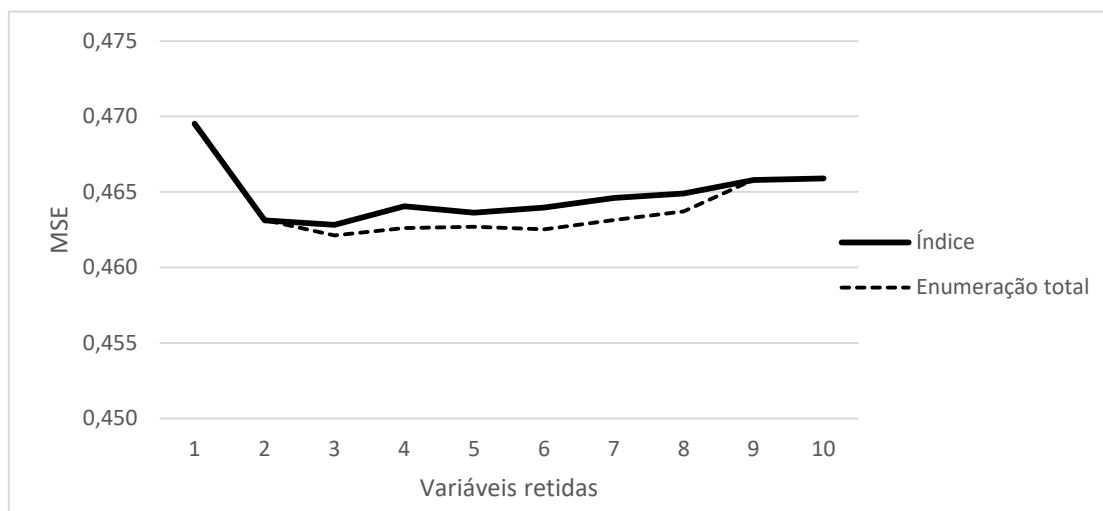


Figura 3.4: Comparação com enumeração total

Ao utilizar a enumeração total, a quantidade ideal de variáveis manteve-se em 3, com um erro de 0,462 (o método selecionou as mesmas duas variáveis de processo e o ácido gálico ao invés do pH). Como este erro é muito semelhante ao obtido pelo índice (0,463), mas o custo computacional da enumeração total é maior, valida-se o índice de seleção de variáveis baseado em elementos do PLS como técnica de seleção de variáveis, principalmente em situações onde a quantidade de variáveis é grande a ponto de tornar a enumeração total inviável.

Embora o erro tenha uma redução pouco significativa (de 0,466 para 0,463), a principal vantagem dessa seleção está na redução dos custos, sendo apenas necessário mensurar o pH da amostra ao invés de serem feitos os demais testes (a classe de altitude e o tipo de fermentação são informações já conhecidas ao serem produzidos os produtos).

Outra maneira de se avaliar a redução do custo a partir da seleção de degustadores e variáveis é comparar a dimensão das matrizes \mathbf{X} e \mathbf{Y} antes e depois da seleção. No estágio inicial, ambas as matrizes continham 396 linhas (9 produtos \times 4 amostras por produto \times 11 degustadores) e 10 colunas (10 variáveis no caso de \mathbf{X} e 10 atributos para \mathbf{Y}). Após a seleção, \mathbf{X} e \mathbf{Y} tiveram o número de linhas reduzido para 180 (9 produtos \times 4 amostras por produto \times 5 degustadores) com 3 colunas no caso de \mathbf{X} (3

variáveis) e 10 colunas para \mathbf{Y} , o que representou uma redução de 67,50% no tamanho total das matrizes.

3.5. Conclusões

O procedimento executado neste artigo consistiu em aplicar um modelo de regressão PLS para prever as notas de um painel sensorial de café, realizando simultaneamente uma seleção de degustadores e de variáveis. Essa seleção foi feita com dois objetivos: (i) reduzir o erro (MSE), tornando a predição mais eficaz, e (ii) reduzir o número de degustadores e variáveis, diminuindo o custo de experimentos futuros e reduzir o erro (MSE).

O primeiro objetivo da seleção foi atingido parcialmente, pois embora tenha havido uma redução no MSE, essa redução foi pequena (de 0,479 para 0,463). No entanto, o segundo objetivo foi atingido plenamente, pois foi possível reduzir o número de degustadores necessários de 10 para 5 e o número de variáveis de 11 para 3. A redução das variáveis foi ainda mais representativa pelo fato das 3 variáveis retidas serem de fácil obtenção, sendo necessário apenas um teste para obter uma variável físico-química (altitude e tipo de fermentação são dados conhecidos de cada amostra).

Ao fazer a seleção de variáveis pelo índice baseado em elementos da regressão PLS, houve uma redução baixa do erro, o que pode indicar que a seleção de variáveis não foi realizada de forma correta ou que a regressão PLS é robusta em relação a quantidade de variáveis utilizadas. No entanto, ao fazer o mesmo procedimento de seleção de variáveis com um método de enumeração total, o erro foi praticamente o mesmo. Com isso, se valida a utilização do índice baseado em PLS (dado que obteve um erro similar ao método de enumeração total com menor custo computacional) e demonstra-se a robustez da regressão PLS.

Por fim, sugere-se para estudos futuros a aplicação das mesmas técnicas para outros bancos de dados, sejam para café ou outros gêneros alimentares. Outra possibilidade é testar outros métodos para prever as notas, como o PCR. Por fim, pode-se usar o PLS com outros índices ou técnicas para a seleção de degustadores ou variáveis.

3.6. Referências

ABDI, H. Partial least squares regression and projection on latent structure regression (PLS Regression). **Wiley Interdisciplinary Reviews: Computational Statistics**, Volume 2, 2010.

ANZANELLO, M.; ALBIN, S.; CHAOVALITWONGSE, W. Selecting the best variables for classifying production batches into two quality levels. **Chemometrics and Intelligent Laboratory Systems**, Volume 97, 2009

ANZANELLO, M.; FOGLIATTO, F.; ROSSINI, K. Data mining-based method for identifying discriminant attributes in sensory profiling. **Food Quality and Preference**, Volume 22, 2011.

BANFIELD, C; HARRIES, J. A technique for comparing judges' performance in sensory tests. **International Journal of Food Science & Technology**, Volume 10, 1975.

BIANCOLILLO, A.; LILAND, K.; MÅGE, I.; NÆS, T.; BRO, R. Variable selection in multi-block regression. **Chemometrics and Intelligent Laboratory Systems**, Volume 156, 2016.

BOLDINA, I.; BENINGER, P. Strengthening statistical usage in marine ecology: Linear regression. **Journal of Experimental Marine Biology and Ecology**, Volume 474, 2016.

BORGOGNONE, M. BUSSI, J.; HOUGH, G. Principal component analysis in sensory analysis: covariance or correlation matrix? **Food Quality and Preference**, Volume 12, 2001.

EFRON, B.; TIBSHIRANI, R. **An introduction to the Bootstrap**. 1. Ed., New York: CRC Press LLC, 1998

GELADI, P.; KOWALSKI, B. Partial least-squares regression: a tutorial. **Analytica Chimica Acta**, Volume 185, 1986.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**: Data mining, Inference and Prediction. 2. Ed., New York: Springer, 2009.

HÖSKULDSSON, A. PLS regression methods. **Journal of chemometrics**, Volume 2, 1988.

JIANG, G.; WANG, W. Error estimation based on variance analysis of k -fold cross-validation. **Pattern Recognition**, Volume 69, 2017.

KIM, J. Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. **Computational Statistics and Data Analysis**, Volume 53, 2009.

KOHAVI, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. **Proceedings of the 14th international joint conference on Artificial Intelligence**, Volume 2, 1995

LATREILLE, J.; MAUGER, E.; AMBROISINE, L.; TENENHAUS, M.; VINCENT, M.; NAVARRO, S.; GUINOT, C. Measurement of the reliability of sensory panel performances. **Food Quality and Preference**, Volume 17, 2006.

LEDAUPHIN, S.; HANAFI, M.; QANNARI, E. Assessment of the agreement among the subjects in fixed vocabulary profiling. **Food Quality and Preference**, Volume 17, 2006.

LENGARD, V.; KERMIT, M. 3-Way and 3-block PLS regressions in consumer preference analysis. **Food Quality and Preference**, Volume 17, 2006.

LOMBARDO, R. The analysis of sensory and chemical-physical variables via multivariate additive PLS splines. **Food Quality and Preference**, Volume 22, 2011.

GODOY, J.; VEGA, J.; MARCHETTI, J. Relationships between PCA and PLS-Regression. **Chemometrics and Intelligent Laboratory Systems**, Volume 130, 2014.

MEHMOOD, T.; LILAND, K.; SNIPEN, L.; SÆBØ, S. A review of variable selection methods in Partial Least Squares Regression. **Chemometrics and Intelligent Laboratory Systems**, Volume 118, 2012.

MONTGOMERY, D.; PECK, E.; VINING, G. Geoffrey. **Introduction to Linear Regression Analysis**. 5. Ed., New Jersey: Wiley, 2012.

POMBEIRO, H.; SANTOS, R.; CARREIRA, P.; SILVA, C.; SOUSA, J. Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks. **Energy and Buildings**, Volume 146, 2017.

ROSSINI, K.; VERDUN, S.; CARIOU, V.; QANNARI, E.; FOGLIATTO, F. PLS discriminant analysis applied to conventional sensory profiling data. **Food Quality and Preference**, Volume 23, 2012.

SCA, 2018. **Specialty Coffee Association Protocols**. 2018. Disponível em [≤https://sca.coffee/research/protocols-best-practices>](https://sca.coffee/research/protocols-best-practices). Acesso em 4 set. 2018.

WANG, Z.; HE, Q.; WANG, J. Comparison of variable selection methods for PLS-based soft sensor modeling. **Journal of Process Control**, Volume 26, 2015.

WILDERJANS, T.; CARIOU, V. CLV3W: A clustering around latent variables approach to detect panel disagreement in three-way conventional sensory profiling data. **Food Quality and Preference**, Volume 47, 2016

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, Volume 58, 2001.

4. Terceiro artigo: Comparação de técnicas de classificação e seleção de variáveis aplicadas a amostras de café descritas por variáveis espectrais

Luiz Antônio Nonenmacher Júnior (ljuniornone@gmail.com)

Michel José Anzanello (anzanello@producao.ufrgs.br)

Resumo

Uma das preocupações recorrentes de produtores do setor alimentício é a qualidade de seus produtos, que frequentemente é mensurada através do uso de painéis sensoriais. Um painel sensorial é composto de degustadores que avaliam as amostras com base em diferentes atributos sensoriais; tais resultados podem ser utilizados, dentre outros propósitos, para dividir os produtos em categorias de qualidade. A partir disso, podem ser utilizadas técnicas para prever a categoria de determinado produto com base em variáveis que o descrevam (variáveis físicas, químicas ou de processo). Neste artigo serão utilizadas técnicas de seleção de variáveis e de classificação para prever a categoria de qualidade de amostras de café especial com base em dados de espectroscopia por infravermelho. Para isso, inicialmente serão utilizados dois índices para seleção de variáveis, sendo o primeiro baseado na análise de componentes principais (PCA) e o segundo na distância de Bhattacharyya. Na sequência, serão testadas três técnicas para classificação: algoritmo do vizinho mais próximo (k -NN), análise linear discriminante (LDA) e rede neural probabilística (PNN). Como resultados, definiu-se o índice baseado em Bhattacharyya como tendo o melhor desempenho, o PNN como a técnica que gerou a melhor acurácia média (0,678) e o LDA como a técnica que reteve o menor percentual médio de variáveis (8,2%).

Palavras-chave: k -NN, LDA, PNN, Seleção de Variáveis, Espectroscopia por infravermelho, Painéis Sensoriais.

Abstract

One of the most recurrent preoccupations of producers in the food industry is related to the quality of its products, which is frequently measured with the utilization of sensory panels. A sensory panel is composed of a group of graders that evaluate the samples in different sensory attributes and its results could be used to divide the products in quality categories. With these categories, several techniques could be used to forecast the

category of some product based on their variables. This study will utilize variable selection and classification techniques to forecast the quality category of special coffee samples using data obtained from an infrared spectroscopy. For this purpose, two indexes for variable selection techniques will be used, based on the principal component analysis (PCA) and the Bhattacharyya distance. After, three classification techniques will be tested: the closest neighbor algorithm (k-NN), the linear discriminant analysis (LDA) and the probabilistic neural network (PNN). As results, the index based in the Bhattacharyya distance presented the best performance, with the PNN being the technique that had the best average accuracy (0.678) and the LDA the technique with the lowest average percentage of variable retention (8.2%).

Keywords: *k-NN*, LDA, PNN, Variables Selection, Infrared Spectroscopy, Sensory Panels.

4.1. Introdução

Uma das preocupações frequentes de produtores do setor alimentício diz respeito à qualidade de seus produtos, visto que ela tem influência direta na demanda e no valor atribuído ao produto. Para mensurar essa qualidade, o setor faz uso de painéis sensoriais (MOSER et al., 2018), os quais consistem em um grupo de degustadores que avaliam amostras de produtos com base em um conjunto de atributos sensoriais. Para obter resultados satisfatórios, um painel deve ser composto por degustadores experientes e extensivamente treinados, e seu desempenho deve ser verificado regularmente (WILDERJANS e CARIOU, 2016). Essas avaliações podem ocorrer ou não com a presença de repetições e os resultados são tipicamente expressos em uma escala numérica contínua. Com base nesses resultados, os produtos podem ser classificados em faixas de atribuição de valor (por exemplo, faixas *premium* e regular).

A fim de ajustar seus processos produtivos, torna-se interessante para os produtores prever a faixa de qualidade de seu produto com base em variáveis físicas, químicas ou de processo que caracterizem o produto. Desta forma, pode-se reduzir a dependência do processo de aferição de qualidade atribuído aos degustadores, os quais tendem a apresentar fadiga em cenários de longas sessões de prova. Além de potencial aumento na precisão de classificação, a utilização de modelos pode ainda reduzir custos atrelados ao painel sensorial.

Neste artigo são comparadas abordagens de seleção de variáveis com vistas à maximização da acurácia de classificação e minimização da quantidade de variáveis retidas para categorização de amostras de alimentos. Os métodos são aplicados em dados espectrais do tipo infravermelho para classificação de amostras de café especial em duas categorias de qualidade (superior e inferior) com relação a sete atributos sensoriais (fragrância, sabor, finalização, acidez, corpo, equilíbrio e final).

A primeira etapa dessa comparação consiste em comparar o desempenho de duas técnicas de seleção de variáveis baseadas em índices de importância. O índice de importância da primeira dessas técnicas é obtido através de uma análise de componentes principais (PCA), enquanto que o índice da segunda técnica é gerado através da distância de Bhattacharyya. Tais índices são integrados a três técnicas de classificação: o algoritmo do vizinho mais próximo (k -NN), a análise linear discriminante (LDA) e a rede neural probabilística (PNN). Os desempenhos das combinações de índices e técnicas de classificação são avaliados em relação ao percentual de variáveis retidas e acurácia de classificação.

O artigo está organizado da seguinte forma: após a introdução, a seção 4.2 apresenta conceitos sobre seleção de variáveis e o problema da classificação, com ênfase às técnicas utilizadas neste artigo. Na sequência, a seção 4.3 apresenta o procedimento utilizado no estudo, detalhando a lógica de seleção de variáveis e classificação. Já a seção 4.4 descreve os resultados obtidos, analisando o desempenho das técnicas de seleção de variáveis e de classificação. A seção 4.5, por fim, arremata o artigo, retomando os objetivos do artigo e sugerindo possibilidades para futuros trabalhos.

4.2. Referencial teórico

Esta seção é subdividida em cinco subseções. A primeira subseção apresenta uma breve introdução sobre técnicas de classificação. Já as subseções 4.2.2, 4.2.3 e 4.2.4 apresentam as técnicas de classificação utilizadas neste artigo: o algoritmo do vizinho mais próximo, a análise discriminante linear e a rede neural probabilística. A seção 4.2.5, por fim, discute seleção de variáveis.

4.2.1 Técnicas de classificação

Técnicas de classificação são utilizadas em situações onde é necessário determinar a classe a que pertence uma determinada amostra de um produto com base em variáveis que a descrevem, sendo uma classe definida como um conjunto de amostras com uma propriedade em comum. No setor alimentício, essa propriedade em comum frequentemente é a autenticidade de um produto, sua região de origem ou faixa de qualidade e preço (CALLAO e RUISÁNCHEZ, 2018).

Todas as amostras de cada classe devem ser descritas pelo mesmo conjunto de variáveis; a partir disso se estabelece uma regra de classificação cujo objetivo é classificar amostras de classe desconhecida. Existem diversas técnicas para obtenção dessa regra de classificação, sendo utilizadas nesse artigo três delas: o algoritmo do vizinho mais próximo (k -NN), a análise discriminante linear (LDA) e a rede neural probabilística (PNN).

Independente da técnica utilizada, é fundamental a utilização de uma métrica para avaliação do desempenho de classificação. Um dos procedimentos frequentemente utilizados consiste em dividir os dados em um banco de treino e um de teste. O modelo de classificação é gerado utilizando os dados do banco de treino e, na sequência, aplica-se o mesmo para se prever as classes das amostras do banco de teste. Dividindo a quantidade de acertos pelo total de amostras obtêm-se a acurácia de classificação.

Além da acurácia de classificação, em algumas situações utiliza-se também uma matriz chamada de matriz de confusão ou matriz de erro. Essa matriz tem o número de colunas e linhas iguais ao número de classes e representa em cada linha as amostras preditas em tal classe e em cada coluna as amostras reais de cada classe (STEHMAN, 1997). No caso de um problema com apenas duas classes (sendo uma classe chamada de positiva e a outra de negativa) os quatro elementos da matriz são chamados de verdadeiros positivos, falsos negativos, falsos positivos e verdadeiros negativos, sendo que as amostras corretamente classificadas encontram-se nos verdadeiros positivos e verdadeiros negativos.

A partir dessa matriz podem ser calculados dois índices: a sensibilidade e a especificidade. A sensibilidade mede a proporção de amostras da classe positiva que são corretamente classificadas como positivas e pode ser calculada através da divisão dos

verdadeiros positivos pela soma dos verdadeiros positivos com os falsos negativos. Já a especificidade mede a proporção de amostras da classe negativa que são corretamente classificadas e pode ser calculada através da divisão dos verdadeiros negativos pela soma dos verdadeiros negativos com os falsos positivos. Dependendo do problema em questão pode-se buscar maximizar um desses índices em detrimento do outro ou atingir um equilíbrio entre ambos (mesmo podendo reduzir a acurácia média do modelo).

4.2.2 Algoritmo do vizinho mais próximo

O algoritmo do vizinho mais próximo, *k*-NN (*k*-Nearest Neighbour), proposto por Fix e Hodges (1951), é uma técnica frequentemente utilizada para problemas de classificação por conta de seus bons resultados e simplicidade. Trata-se de uma técnica supervisionada e não paramétrica que consiste em, dado um ponto δ cuja classe é desconhecida, calcular a distância desse ponto em relação aos demais pontos cuja classe é conhecida. Selecionam-se então os *k* pontos mais próximos e faz-se uma contagem da classe de cada um, sendo a classe mais frequente atribuída a δ (BAJOUB et al., 2017; RENCHER, 2002).

k é o parâmetro mais importante do método e costuma-se escolher valores pequenos (de 1 a 5) para obtenção de melhores resultados (BAJOUB et al., 2017). Para a definição do *k* a ser utilizado, recomenda-se testar uma faixa de valores e selecionar aquele que apresenta maior capacidade preditiva. O *k*-NN pode ser implementado com diferentes medidas de distância, sendo utilizada nesse artigo a distância Euclidiana.

A principal desvantagem do *k*-NN é sua sensibilidade à estrutura local dos dados, sendo sua capacidade preditiva reduzida em situações onde pontos de mesma classe estejam distantes entre si no espaço (RASTEGARZADEH e NEMATI, 2018). Além disso, embora seja um algoritmo simples conceitualmente, sua execução é lenta para um número elevado de variáveis, quando comparado a outras técnicas de classificação.

No setor alimentício, diversos artigos utilizam o *k*-NN como ferramenta de classificação. Bajoub et al. (2017) aplicaram a técnica para classificar geograficamente amostras de azeite de oliva. Já Santos et al. (2016) fizeram uso do *k*-NN para avaliar e controlar a adulteração em amostras de leite. Por fim, o *k*-NN foi utilizado por Drivelos et al. (2014) para discriminar espécie de ervilhas.

4.2.3 Análise discriminante linear

A análise discriminante linear, conhecida como LDA (*linear discriminant analysis*), é uma técnica introduzida por Fisher (1936) que pode ser utilizada para fins de classificação ou de redução de dimensionalidade de um conjunto de dados. O LDA determina um novo sistema de coordenadas que maximiza a separação entre as categorias, sendo a primeira dimensão definida como aquela que maximiza a razão entre a variância interclasse e intraclasse (FORINA et al., 2015; FUKUNAGA, 1990). Esse novo sistema de coordenadas pode então ser utilizado para classificar as novas amostras.

A análise discriminante linear é similar à Análise de Componentes Principais (PCA) no sentido de que ambas são transformações lineares que buscam alterar o sistema de coordenadas para maximizar a variância. No entanto, a principal diferença entre as duas técnicas é que o PCA é um algoritmo não supervisionado, que busca maximizar a variância entre todos os dados, enquanto o LDA é um algoritmo supervisionado que busca maximizar a variância entre as classes (MARTÍNEZ e KAK, 2001).

A principal limitação do LDA refere-se à relação entre a quantidade de amostras e de variáveis. É recomendado que sejam utilizadas mais amostras do que variáveis (idealmente, o número de amostras é de três a cinco vezes o número de variáveis) para se obter uma capacidade preditiva confiável (ALAMPRESE et al., 2013). Essa necessidade pode se tornar uma limitação em problemas cujas amostras são descritas por um grande conjunto de variáveis, sendo frequentemente utilizadas, nestas situações, técnicas para selecionar e reduzir a quantidade de variáveis (FORINA et al., 2015).

O LDA é uma técnica comumente utilizada no ramo alimentício para classificação. O LDA foi utilizado por Forina et al. (2015) para separar amostras de azeite de oliva provenientes de uma região específica da Itália de outras amostras vindas de outros locais do país. A técnica foi utilizada também por Erich et al. (2015) para a diferenciação entre leite orgânico e convencional. Já Alamprese et al. (2013) utilizaram o LDA para detectar alterações em amostras de carne moída de peru.

4.2.4 Redes neurais probabilísticas

Redes neurais artificiais são sistemas matemáticos inspirados nas redes neurais biológicas encontradas nos cérebros de animais. Essas redes são constituídas de um conjunto de neurônios artificiais distribuídos em camadas, sendo uma rede geralmente constituída de uma camada de entrada, camadas intermediárias e uma camada de saída. Cada neurônio recebe sinais de outros neurônios, os processa e os envia para outros neurônios, até gerar um valor de saída da rede. Tipicamente o sinal entre neurônios é um número real e a saída de cada neurônio é obtida por uma função não-linear de soma de suas entradas (PRIETO et al., 2016; RIPLEY, 1996).

Redes neurais têm sido utilizadas com os mais diversos fins, incluindo previsão de séries temporais, classificação e reconhecimento de fala, escrita e imagens (CAO et al., 2018). A utilização de uma rede neural inicia-se pela alimentação de dados, seguida da fase de treinamento, onde os parâmetros (pesos) de cada ligação são ajustados para se obter o melhor resultado. Existem diversos algoritmos para esse treinamento, sendo o *back propagation* o mais tradicional e consistindo em pequenas mudanças nos pesos que gradualmente melhoram o resultado da rede. No entanto, esse algoritmo requer um grande esforço computacional, além de estar sujeito a mínimos locais (SPECHT, 1989). Como forma de contornar esses problemas, Specht (1989) propôs as redes neurais probabilísticas (PNN, *probabilistic neural network*).

Uma rede neural probabilística é uma rede de tipo *feedforward*, adaptada para problemas de classificação e baseada em redes bayesianas. Ela é composta de quatro camadas: camada de entrada, camada oculta, camada de soma e camada de saída (RASTEGARZADEH e NEMATI, 2018). A primeira camada recebe um determinado vetor de entrada a ser classificado e a camada oculta calcula a distância desse vetor para o vetor médio de cada classe (obtido através da média dos vetores de treino de cada classe). A camada de soma então gera um conjunto de probabilidades que representa a chance desse vetor pertencer a cada classe e, por fim, a última camada seleciona a maior probabilidade e atribui o vetor de entrada à essa classe (SPECHT, 1989).

4.2.5 Seleção de variáveis

Uma característica presente em muitos problemas de classificação é a grande quantidade de variáveis, que geralmente excede o número de amostras. Além de

tipicamente reduzir o desempenho da classificação, uma quantidade elevada de variáveis impede o uso de algumas técnicas estatísticas (WANG et al., 2018). Para contornar essa limitação, é frequente a presença de uma etapa anterior à classificação com vistas à redução do número de variáveis.

Existem duas abordagens para diminuição da quantidade de variáveis: redução de dimensão e seleção de variáveis. A redução de dimensão consiste em encontrar um novo espaço com menor dimensão para substituir o espaço atual das amostras, o que pode ser obtido com técnicas como o PCA (*Principal component analysis*) e o PLS (*Partial least squares*) (WANG et al., 2018). Já a seleção de variáveis consiste em identificar um subconjunto das variáveis originais que contenham a parte mais relevante da informação encontrada no conjunto original de variáveis (PERES e FOGLIATTO, 2018).

Existem diversas vantagens de se aplicar a seleção de variáveis. Uma delas é que modelos que utilizem uma grande quantidade de variáveis tendem a gerar *overfitting*, fazendo com que o ruído dos dados seja confundido com as relações reais, diminuindo a capacidade preditiva do mesmo para dados futuros (VALBUENA et al., 2017). Além disso, um modelo com um menor número de variáveis é mais fácil de ser interpretado e a coleta dos dados para sua execução futura é facilitada (ANZANELLO e FOGLIATTO, 2014).

Existem diferentes técnicas para a seleção de variáveis, como a seleção para frente (*forward selection*), seleção para trás (*backward selection*), PLS, PCA, clusterização e outras ferramentas de mineração de dados (PERES e FOGLIATTO, 2018). Uma dessas abordagens, que será utilizada neste estudo, consiste em atribuir a cada variável um índice de desempenho que represente o grau de importância daquela variável frente às demais (obtido através de técnicas como PCA, PLS, clusterização ou cálculo de distâncias). Na sequência, inicia-se o modelo com todas as variáveis e retira-se, a cada passo, a variável de menor importância.

Diversos artigos no setor de alimentos utilizam a seleção de variáveis antes de aplicar técnicas de classificação ou previsão. Miaw et al. (2018) utilizaram o iPLS (*interval partial least squares*) como algoritmo de seleção de variáveis na detecção de adulteração em néctares de frutas, permanecendo com 7,0% das variáveis originais e

reduzindo o erro de previsão. Já a ANOVA (*Analysis of variance*) foi utilizada como técnica de seleção de variáveis no trabalho de Sherman et al. (2018), que correlacionou a composição volátil de vinhos com dados sensoriais do mesmo.

4.3. Procedimentos metodológicos

O objetivo deste artigo é comparar o desempenho de três técnicas de classificação e dois índices de seleção de variáveis na classificação de amostras de café especial processados por via-úmida em duas categorias de qualidade (superior e inferior) para sete atributos sensoriais (fragrância, sabor, finalização, acidez, corpo, equilíbrio e final). As classes de cada amostra em cada atributo foram obtidas através dos resultados de um painel sensorial, que seguiu o protocolo da *Specialty Coffee Association of America (SCA, 2018)* e foi aplicado por degustadores certificados pelo CQI (*Coffee Quality Institute*). Já as variáveis que descrevem cada amostra são o resultado de uma análise de espectroscopia do tipo infravermelho.

Os dados utilizados neste artigo são representados por duas matrizes: \mathbf{X} e \mathbf{Y} . A matriz \mathbf{X} é composta de N linhas e J colunas, onde N é a quantidade de amostras (999), J é o total de variáveis geradas pela análise de espectroscopia (3596) e \mathbf{X}_{nj} representa o valor da variável j para a amostra n . \mathbf{Y} é composto de N linhas e M colunas, sendo M a quantidade de atributos avaliados (7). Cada elemento de \mathbf{Y} , \mathbf{Y}_{nm} representa a classe da amostra n (superior ou inferior) para o atributo avaliado m .

A primeira etapa consiste em processar os dados da espectroscopia (matriz \mathbf{X}) pelo método SNV (*Standard normal variate*), corrigindo os efeitos de dispersão de luz. Esse método consiste em calcular a média e o desvio padrão de cada variável de \mathbf{X} . Na sequência, cada elemento é subtraído da média de sua variável e dividido pelo desvio padrão da mesma, de tal forma que as médias de cada variável se tornam zero e os desvios 1.

Na sequência, são calculados os índices de importância para cada variável. São calculados dois índices, do tipo maior-é-melhor, sendo o índice 1 baseado em coeficientes do PCA e o índice 2 na distância de Bhattacharyya.

O primeiro índice prioriza as variáveis que explicam uma parcela maior da variância encontrada nos dados através de elementos obtidos pela aplicação do PCA. A

aplicação da análise de componentes principais gera as variáveis latentes t a partir de combinações lineares das variáveis originais de \mathbf{X} associadas a pesos w , conforme equação (1), na qual a varia de 1 até o número de variáveis J (WOLD et al., 2001).

$$t_{na} = w_{1a}x_{n1} + w_{2a}x_{n2} + \dots + w_{Ja}x_{nJ} = \mathbf{W}'\mathbf{X} \quad (1)$$

Além disso, dividindo a variância de cada variável latente pela variância total, obtêm-se o percentual de variância explicado por cada variável, $p = (p_1, p_2, \dots, p_J)$. O índice de importância de cada variável j , v_j é obtido pela multiplicação de p pelo módulo do peso w correspondente, conforme equação (2):

$$v_j = p_1|w_{j1}| + p_2|w_{j2}| + \dots + p_J|w_{jJ}| \quad (2)$$

A principal limitação do primeiro índice é que o mesmo é calculado analisando apenas a distribuição da matriz \mathbf{X} , desconsiderando a classe de cada amostra. Já o segundo índice proposto não apresenta tal limitação pois avalia o quanto cada variável é diferente ao descrever amostras de classes distintas. Para isso, cada variável é decomposta em dois conjuntos (g e h) segundo as classes das amostras e calcula-se a diferença entre esses dois conjuntos através da distância de Bhattacharyya. Assumindo duas classes normalmente distribuídas, Coleman e Andrews (1975), propuseram uma fórmula de cálculo que utiliza apenas as médias (μ_g e μ_h) e as variâncias (σ_g^2 e σ_h^2) de cada classe, conforme a equação 3:

$$DB = \frac{1}{4} \ln \left(\frac{1}{4} \left(\frac{\sigma_g^2}{\sigma_h^2} + \frac{\sigma_h^2}{\sigma_g^2} + 2 \right) \right) + \frac{1}{4} \left(\frac{(\mu_g - \mu_h)^2}{\sigma_g^2 + \sigma_h^2} \right) \quad (3)$$

Pela equação 3 é possível perceber que essa distância é diretamente proporcional à diferença de médias entre os conjuntos e inversamente proporcional ao tamanho das variâncias. Além disso o primeiro termo da equação cresce com a diferença entre os desvios dos conjuntos, de tal forma que em uma situação onde as médias são iguais, dois conjuntos de variâncias diferentes terão uma distância maior que dois conjuntos de mesma variância.

Após o cálculo dos índices, o próximo passo consiste em executar os experimentos para cada combinação de índices (1 e 2), técnicas (k -NN, LDA e PNN) e atributos, gerando um total de 42 experimentos (2 índices, 3 técnicas e 7 atributos). Para execução do k -NN utilizou-se um valor de k igual a 1 e para o PNN utilizou-se o

parâmetro sigma igual a 0,1, definidos através de procedimentos de validação cruzada (RENCHEER, 2002).

Cada experimento é executado segundo o seguinte procedimento: aplica-se a técnica de classificação sobre todo o conjunto de variáveis e coleta-se a acurácia de classificação. Na sequência, retira-se a variável de menor importância (menor valor do índice) e executa-se novamente a classificação, coletando a nova acurácia. Esse procedimento segue até que reste apenas uma variável, gerando uma tabela que relaciona o número de variáveis utilizadas à acurácia de classificação. Esse mesmo procedimento é repetido cem vezes e, após a centésima iteração, obtêm-se uma acurácia e um percentual de variáveis retidas através da média da maior acurácia e percentual de variáveis retidas correspondente de cada uma das cem iterações. A utilização de repetições assegura que o método não é beneficiado ou penalizado por uma única partição nas porções de treino e teste.

De posse desses resultados, são realizadas as comparações. A primeira comparação é entre os índices, visando a identificar o índice que maximiza a acurácia e minimiza o percentual de variáveis retidas. Na sequência, analisam-se as técnicas de classificação, identificando a responsável pela maior acurácia e a menor quantidade de variáveis retidas. A partir dessas duas comparações é definida uma combinação sugerida de índice e técnica para ser usado em futuras classificações usando dados semelhantes. Por fim, faz-se uma análise dos resultados de cada atributo.

4.4. Resultados

Essa seção está dividida em três etapas, sendo na primeira avaliado o desempenho dos índices, na segunda o desempenho das técnicas e na terceira os sete atributos sensoriais.

4.4.1 Análise dos índices

Para realizar a escolha do melhor índice, comparou-se a acurácia e percentual de variáveis retidas para cada combinação de índice e técnica de classificação, calculando a média desses indicadores nos sete atributos sensoriais. O resultado dessa comparação é apresentado na Tabela 4.1.

Tabela 4.1: Acurácia e % de variáveis retidas para os 2 índices de importância

	Acurácia		% de variáveis retidas	
	Índice 1	Índice 2	Índice 1	Índice 2
k-NN	0,622	0,656	60,30%	19,50%
LDA	0,615	0,662	88,40%	8,20%
PNN	0,635	0,678	65,70%	11,90%
Média	0,624	0,665	71,5%	13,2%

Como pode ser observado na Tabela 4.1, o índice 2 (baseado na distância de Bhattacharyya) gerou uma acurácia média 6,6% maior que o índice 1, baseado em PCA. No entanto, a maior distinção entre os índices encontrou-se na quantidade de variáveis retidas. Enquanto o índice 1 reteve em média 2569 das 3596 variáveis, o segundo índice reteve apenas 476 variáveis, o que torna o modelo mais robusto frente à *overfitting*. Outro resultado importante é que o índice 2 teve um desempenho melhor para as três técnicas de classificação, tanto em acurácia quanto em variáveis retidas, o que corrobora sua utilização nas próximas etapas da análise.

Embora o PCA utilize mais parâmetros em seu cálculo que a distância de Bhattacharyya (o que tende a gerar melhores resultados na seleção de variáveis com alta correlação), o índice baseado no PCA teve um pior desempenho. Isso pode ser explicado pelo funcionamento desse índice, que é calculado desconsiderando a classe de cada amostra (ele avalia somente as relações de variabilidade dentre as variáveis espectrais). Dessa forma, uma variável que apresentasse grande variabilidade dentro de uma mesma classe ao invés de para classes diferentes conduziria a um elevado índice de importância, piorando o desempenho de seleção.

4.4.2 Análise das técnicas de classificação

Conforme resultado em 4.4.1, recomenda-se a utilização do índice 2 para ranqueamento das variáveis de acordo com sua relevância para classificação das amostras de café. Com base neste índice, calculou-se a acurácia e percentual de variáveis retidas para cada combinação de atributos sensoriais e técnicas, conforme apresentado na Tabela 4.2.

Tabela 4. 4.2: Acurácia e % de variáveis retidas para as 3 técnicas de classificação

Atributo	Acurácia			% de variáveis retidas		
	k-NN	LDA	PNN	k-NN	LDA	PNN
Fragrância	0,649	0,643	0,674	3,9%	6,3%	8,6%
Sabor	0,656	0,678	0,692	14,1%	2,7%	5,6%
Finalização	0,671	0,664	0,692	44,7%	8,4%	10,9%
Acidez	0,669	0,664	0,681	22,6%	18,3%	19,3%
Corpo	0,636	0,652	0,657	5,3%	10,5%	9,6%
Equilíbrio	0,649	0,660	0,670	39,1%	7,5%	19,5%
Final	0,659	0,670	0,681	7,3%	4,0%	9,9%
Média	0,656	0,662	0,678	19,6%	8,2%	11,9%

De acordo com esses resultados, percebe-se que o *k*-NN apresentou o pior desempenho dentre as técnicas testadas, conduzindo à menor acurácia e maior percentual médio de variáveis retidas. Esse menor desempenho pode ser explicado pelas limitações do *k*-NN, que é uma dependente da localização das amostras no espaço, sendo necessário para um bom desempenho que as amostras de mesma classe estejam próximas entre si, o que não ocorreu com os dados analisados. Os dados do *k*-NN mostrados na tabela são referentes a um valor de *k* igual a 1, mas testes preliminares com maiores valores de *k* geraram resultados piores em termos de acurácia e percentual de variáveis retidas.

Já o LDA apresentou uma acurácia melhor que o *k*-NN, mas pior que o PNN, gerando, no entanto, o menor percentual de variáveis retidas. Esse comportamento está relacionado à característica do LDA de ter um desempenho superior quando o número de variáveis é menor do que o número de amostras (idealmente de 3 a 5 vezes); no caso em estudo, o número médio de variáveis retidas foi 295, o que é 3,38 vezes menor do que as 999 amostras.

Outra forma de avaliar a relação entre variáveis retidas e acurácia consiste em analisar a relação entre esses dois indicadores para cada atributo avaliado pelo LDA. De acordo com os resultados da Tabela 4.2, percebe-se que os dois melhores resultados em termos de acurácia foram obtidos com os dois menores percentuais de variáveis retidas (Sabor e Final). Além disso, nota-se que os dois maiores percentuais de variáveis retidas (Acidez e Corpo) geraram o segundo e quarto piores resultados em termos de acurácia.

Por fim, o PNN apresentou a melhor acurácia das três técnicas de classificação e o segundo menor percentual de variáveis retidas. No que se refere a acurácia, é importante notar que, além de gerar a melhor acurácia média, o PNN obteve também a melhor acurácia para cada um dos atributos avaliados, demonstrando a robustez e consistência da técnica.

Com relação à quantidade de variáveis retidas, o PNN apresentou um comportamento consistente em todos os atributos. Se analisados os sete atributos, em cinco desses (Sabor, Finalização, Acidez, Corpo e Equilíbrio) o PNN foi o segundo método em termos de retenção de variáveis, sendo o pior nos dois restantes (Fragrância e Final). Uma forma alternativa de analisar a consistência do PNN é calcular, para cada técnica, o coeficiente de variação para o % de variáveis retidas (dividindo a média pelo desvio padrão), conforme Tabela 4.3.

Tabela 4.3: Coeficiente de variação de cada método

	k-NN	LDA	PNN
Média	19,50%	8,20%	11,90%
Desvio Padrão	0,153	0,048	0,05
Coeficiente de Variação	78,5%	58,5%	42,0%

Através do coeficiente de variação, pode-se perceber que o PNN é a técnica mais estável, seguido pelo LDA e *k*-NN. Percebe-se ainda que, para nenhum dos sete atributos, o PNN foi a técnica o que apresentou a menor retenção de variáveis. Calculando a retenção mínima para cada método (considerando os sete atributos), o PNN apresentou o maior valor (5,6% contra 3,9% do *k*-NN e 2,7% do LDA). Isso demonstra que o PNN se comportou de forma robusta frente a uma maior quantidade de variáveis (ao contrário do LDA); por outro lado, o método consistentemente demandou um número maior de variáveis para gerar o melhor resultado em termos de acurácia.

Por fim, a escolha da técnica a ser utilizada depende dos objetivos específicos da análise. Caso reduzir o número de variáveis seja mais importante do que gerar classificações acuradas, deve-se escolher o LDA; caso a acurácia tenha mais importância, deve-se priorizar o PNN. No caso deste artigo, para a análise seguinte (análise dos atributos) escolheu-se priorizar a acurácia, selecionando como base o PNN.

4.4.3 Análise dos atributos

Para analisar os atributos, montou-se a Tabela 4.4 (que é um resumo da Tabela 4.2) mostrando apenas as colunas de acurácia e percentual de variáveis selecionadas utilizando o PNN.

Tabela 4.4: Dados por atributo

Atributo	Acurácia	% de variáveis retidas
Sabor	0,692	5,6%
Finalização	0,692	10,9%
Acidez	0,681	19,3%
Final	0,681	9,9%
Fragrância	0,674	8,6%
Equilíbrio	0,670	19,5%
Corpo	0,657	9,6%
Média	0,678	11,9%

Pela tabela, pode-se perceber que há uma diferença em termos de acurácia entre os atributos, principalmente para o corpo, que apresenta a menor acurácia. Essa menor acurácia pode ser um indicativo de que esse atributo tem um caráter subjetivo mais elevado e que há divergências por parte do painel sensorial na sua mensuração.

Além disso, percebe-se também uma diferença no porcentual de variáveis retidas, com o sabor retendo apenas 5,6% das variáveis enquanto o equilíbrio e a acidez retêm mais de 19%. Essa diferença de retenção está relacionada à quantidade de intervalos de frequência utilizados para se classificar tal atributo; atributos que necessitam de mais variáveis podem representar características sensoriais mais complexas em termos químicos.

4.5. Conclusões

Este artigo utilizou amostras de café para comparar o desempenho de distintos índices de importância e técnicas de classificação para categorizar tais amostras em duas classes de qualidade de acordo com sete atributos sensoriais (através do uso de painéis sensoriais). Os objetivos desse artigo foram atingidos de forma parcial. No que se refere à comparação das técnicas, o objetivo foi alcançado, sendo o índice de importância baseado em distância de Bhattacharyya selecionado como o melhor índice e o PNN e LDA selecionados como melhores técnicas, dependendo do objetivo da análise (maximizar acurácia ou minimizar a quantidade de variáveis retidas). Além disso, o

objetivo de reduzir o número de variáveis utilizadas foi atingido, dado que se conseguiu obter os melhores resultados de acurácia utilizando em média 11,9% (PNN) ou 8,2% (LDA) das variáveis originais.

No entanto, o objetivo de alcançar uma elevada acurácia não foi atingido, dado que a maior acurácia média obtida foi de 0,678. Como sugestão de estudos futuros, a fim de entender essa acurácia, pode-se aplicar as mesmas técnicas deste estudo para outros bancos de dados de café ou outros alimentos, verificando se a acurácia obtida apresenta variação em relação aos resultados obtidos neste estudo.

Outro desdobramento para estudos futuros consiste em analisar as variáveis retidas para cada atributo, buscando relações no espectro gerado que vinculem os atributos sensoriais à constituição química das amostras. A partir disso podem ser geradas sugestões de mudanças que os produtores devem fazer em seus processos para melhorar a qualidade de seus produtos. Por fim, sugere-se também aplicar o mesmo conjunto de técnicas de seleção de variáveis e classificação em bancos compostos por outro tipo de variáveis, avaliando se as mesmas técnicas obtêm um bom desempenho.

4.6. Referências

ALAMPRESE, C.; CASALE, M.; SINELLI, N.; LANTERI, S.; CASIRAGHI, E. Detection of minced beef adulteration with turkey meat by UV-vis, NIR and MIR spectroscopy. **LWT – Food Science and Technology**, Volume 53, 2013.

ANZANELLO, M.; FOGLIATTO, F. A review of recent variable selection methods in industrial and chemometrics applications. **European J. of Industrial Engineering**, Volume 8, 2014.

BAJOUB, A.; MEDINA-RODRÍGUEZ, S.; GÓMEZ-ROMERO, M.; AJAL, E.; BAGUR-GONZÁLEZ, M.; FERNÁNDEZ-GUTIÉRREZ, A.; CARRASCO-PANCORBO, A. Assessing the varietal origin of extra-virgin olive oil using liquid chromatography fingerprints of phenolic compound, data fusion and chemometrics. **Food Chemistry**, Volume 215, 2017.

CALLAO, M.; RUISÁNCHEZ, I. An overview of multivariate qualitative methods for food fraud detection. **Food Control**, Volume 86, 2018.

CAO, W.; WANG, X.; MING, Z.; GAO, J. A review on neural networks with random weights. **Neurocomputing**, Volume 275, 2018.

COLEMAN, G.; ANDREWS, H. Image segmentation by clustering. **Proceedings of the IEEE**, Volume 67, 1979.

DRIVELOS, S.; HIGGINS, K.; KALIVAS, J.; HAROUTOUNIAN, S.; GEORGIU, C. Data fusion for food authentication. Combining rare earth elements and trace metal to discriminate “Fava Santorinis” from other yellow split peas using chemometric tools. **Food Chemistry**, Volume 165, 2014.

ERICH, S.; SCHILL, S.; ANNWEILER, E.; WAIBLINGER, H.; KUBALLA, T.; LACHENMEIER, D.; MONAKHOVA, Y. Combined chemometric analysis of ^1H NMR, ^{13}C NMR and stable isotope data to differentiate organic and conventional milk. **Food Chemistry**, Volume 188, 2015.

FISHER, R. The use of multiple measurements in taxonomic problems. **Annals of human genetics**, Volume 2, 1936.

FIX, E.; HODGES, J. Discriminatory Analysis. Nonparametric discrimination: small sample performance. **USAF School of Aviation Medicine**, Randolph Field, Texas, 1952.

FORINA, M.; OLIVERI, P.; BAGNASCO, L.; SIMONETTI, R.; CASOLINO, M.; GRIFI, F.; CASALE, M. Artificial nose, NIR and UV-visible spectroscopy for the characterization of the PDO Chianti Classico olive oil. **Talanta**, Volume 144, 2015.

FUKUNAGA, K. **Introduction to Statistical Pattern Recognition**, 2nd ed., Academic Press, 1990.

MARTÍNEZ, A.; KAK, A. PCA versus LDA. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Volume 23, 2001.

MIAW, C.; ASSIS, C.; SILVA, A.; CUNHA, M.; SENA, M.; DE SOUZA, S. Determination of main fruits in adulterated nectars by ATR-FTIR spectroscopy combined with multivariate calibration and variable selection methods. **Food chemistry**, Volume 254, 2018.

MOSER, M.; LEPAGE, M.; PINEAU, N.; FILLION, L.; RYTZ, A. Replicates in sensory profiling: Quantification of the impact of moving from two to one assessments. **Food Quality and Preference**, Volume 65, 2018.

PERES, F.; FOGLIATTO, F. Variable selection methods in multivariate statistical process control: A systematic literature review. **Computers & Industrial Engineering**, Volume 115, 2018.

PRIETO, A.; PRIETO, B.; ORTIGOSA, E.; ROS, E.; PELAYO, F.; ORTEGA, J.; ROJAS, I. Neural networks: An overview of early research, current frameworks and new challenges. **Neurocomputing**, Volume 214, 2016.

RASTEGARZADEH, G.; NEMATI, M. Primary mass discrimination of high energy cosmic rays using PNN and *k*-NN methods. **Advances in Space Research**, Volume 61, 2018.

RENCHER, A. **Methods of Multivariate Analysis**, 2nd Ed., John Wiley & Sons, 2002.

RIPLEY, B. **Pattern Recognition and Neural Networks**, Cambridge: Cambridge University Press, 1996.

SANTOS, P.; PEREIRA-FILHO, E.; COLNAGO, L. Detection and quantification of milk adulteration using time domain nuclear magnetic resonance (TD-NMR). **Microchemical Journal**, Volume 124, 2016.

SCA, 2018. **Specialty Coffee Association Protocols**. 2018. Disponível em [≤https://sca.coffee/research/protocols-best-practices>](https://sca.coffee/research/protocols-best-practices). Acesso em 4 set. 2018.

SHERMAN, E.; HARBERTSON, J.; GREENWOOD, D.; VILLAS-BÔAS, S.; FIEHN, O.; HEYMANN, H. Reference samples guide variable selection for correlation of wine sensory and volatile profiling data. **Food Chemistry**, 2018.

SPECHT, D. Probabilistic Neural Networks. **Neural Networks**, Volume 3, 1990.

STEHMAN, S. Selecting and interpreting measures of thematic classification accuracy. **Remote Sensing of Environment**, Volume 62, 1997.

VALBUENA, R.; HERNANDO, A.; MANZANERA, J.; GÖRGENS, E.; ALMEIDA, D.; MAURO, F.; GARCÍA-ABRIL, A.; COOMES, D. Enhancing of accuracy assessment for forest above-ground biomass estimates obtained from remote sensing via hypothesis testing and overfitting evaluation. **Ecological Modelling**, Volume 366, 2017.

WANG, L.; LIN, Y.; WANG, X.; XIAO, N.; XU, Y.; LI, H.; XIU, Q. A selective review and comparison for interval variable selection in spectroscopic modeling. **Chemometrics and Intelligent Laboratory Systems**, Volume 172, 2018.

WILDERJANS, T.; CARIOU, V. CLV3W: A clustering around latent variables approach to detect panel disagreement in three-way conventional sensory profiling data. **Food Quality and Preference**, Volume 47, 2016

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, Volume 58, 2001.

ZHANG, Q.; XIA, D.; WANG, G. Three-way decision model with two types of classification errors. **Information Sciences**, Volume 420, 2017.

5. Considerações finais

Este capítulo apresenta as conclusões da dissertação e sugestões para trabalhos futuros.

5.1 Conclusões

O objetivo principal dessa dissertação foi aplicar técnicas multivariadas sobre dados de painéis sensoriais de café com vistas a mensurar o desempenho de degustadores e prever notas/classes das amostras. Esse objetivo foi decomposto em quatro objetivos específicos: (i) comparar o desempenho de três métodos de treinamento de degustadores de um painel sensorial, encontrando o método que se aplica melhor a cada situação, (ii) aplicar a regressão PLS para prever notas conferidas por degustadores a amostras de café em relação a dez atributos sensoriais, utilizando variáveis de processo e físico-químicas, (iii) utilizar o Índice Alfa de Ledauphin e um índice baseado em PLS para fazer a seleção de degustadores e variáveis a fim de aumentar a capacidade preditiva da regressão PLS e reduzir o custo de futuros experimentos, e (iv) comparar o desempenho de dois índices de seleção de variáveis (baseados em PCA e distância de Bhattacharyya) e três técnicas de classificação (k -NN, LDA e PNN) utilizados para a classificação de amostras de café em duas categorias de qualidade, com base em dados espectrais.

O objetivo (i) foi alcançado no primeiro artigo, no qual foram testados três métodos de treinamento (treinamento de atributos, treinamento de degustadores e treinamento combinado), simulando os resultados através de uma matriz ponderada de notas e de um índice de aprendizado. Após a aplicação dos métodos, definiu-se o treinamento combinado como o que apresenta melhor desempenho para todos os valores de índice de treinamento testados, apresentando uma necessidade de treinamento, em média, 18,7% menor que o treinamento por degustador e 51,4% menor que o treinamento por atributos.

Os objetivos (ii) e (iii) foram atingidos pelo segundo artigo. A aplicação inicial da regressão PLS utilizou um componente, 11 degustadores e 10 variáveis independentes, gerando um erro quadrado médio (MSE) de 0,479. Com a utilização do Índice Alfa de Ledauphin, foram selecionados cinco dos onze degustadores, reduzindo o MSE para 0,466. Por fim, com uso de um índice baseado em parâmetros do PLS, foram retidas três das dez variáveis originais (altitude, pH e tipo de fermentação), obtendo um

MSE de 0,463. A principal contribuição da seleção de degustadores e variáveis foi a redução do custo de futuros experimentos, além de permitir aos produtores identificar as variáveis de processo e físico-químicas que mais contribuem para a qualidade do café.

O objetivo (iv) foi atingido no terceiro artigo, onde foram testadas as combinações de três técnicas de classificação (k -NN, LDA e PNN) e dois índices de importância de variáveis (baseados em distância de Bhattacharyya e PCA) para prever a classificação das amostras em duas categorias de qualidade. No que se refere aos índices, o índice baseado na distância de Bhattacharyya apresentou melhor desempenho para as três técnicas de classificação, tanto em termos de acurácia quanto em quantidade de variáveis retidas. Em relação às técnicas, o PNN apresentou a maior acurácia média (0,678), enquanto que o LDA reteve a menor quantidade de variáveis (8,2%).

As técnicas multivariadas utilizadas nessa dissertação abordaram diferentes usos de painéis sensoriais, sendo utilizadas para treinar degustadores, prever notas conferidas por degustadores e classificar os produtos em categorias de qualidade. Essas aplicações permitem reduzir custos com treinamento e com futuros experimentos, além de fornecer informação que poderá ser utilizada para aumentar a qualidade dos produtos.

5.2 Sugestões para trabalhos futuros

Como extensões das técnicas aplicadas nesta dissertação, são sugeridas as seguintes possibilidades para pesquisas futuras:

- Utilizar os métodos de treinamento apresentados em situações práticas, avaliando se seu desempenho é similar ao obtido nesta dissertação. Além disso, verificar se a simulação do treinamento utilizando uma matriz ponderada de notas e um índice de aprendizado é consistente com a realidade.
- Testar outras técnicas de previsão para prever as notas conferidas aos produtos com base em variáveis de processo e físico-químicas, e comparar os resultados obtidos com os apresentados pela regressão PLS nesta dissertação;
- Realizar uma análise das variáveis espectrais retidas na classificação dos produtos em categorias de qualidade, buscando relações químicas que justifiquem as diferenças de classificação; e

- Aplicar as mesmas técnicas multivariadas apresentadas nesta dissertação a outras amostras de café ou gêneros alimentícios, avaliando se os resultados apresentam variação em relação aos resultados obtidos pela dissertação;