



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS, TECNOLOGIAS E SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS DA INFORMAÇÃO E
COMUNICAÇÃO

Felipe Veira Roque

**Um Framework Conceitual para análise e comparação de padrões em
apresentações orais de estudantes**

Araranguá
2020

Felipe Veira Roque

Um Framework Conceitual para análise e comparação de padrões em apresentações orais de estudantes

Dissertação submetida ao Programa de Pós-Graduação em Tecnologias da Informação e Comunicação da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Tecnologias da Informação e Comunicação.

Orientador: Prof. Cristian Cechinel, Dr.

Coorientador: Prof. Robson Lemos, Dr.

Araranguá

2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Roque, Felipe Vieira

Um Framework Conceitual para análise e comparação de padrões em apresentações orais de estudantes / Felipe Vieira Roque ; orientador, Cristian Cechinel, coorientador, Robson Lemos, 2020.

79 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Campus Araranguá, Programa de Pós-Graduação em Tecnologias da Informação e Comunicação, Araranguá, 2020.

Inclui referências.

1. Tecnologias da Informação e Comunicação. 2. Multimodal Learning Analytics. 3. Apresentações Oraís. 4. Mineração de Dados Educacionais. I. Cechinel, Cristian. II. Lemos, Robson. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Tecnologias da Informação e Comunicação. IV. Título.

Felipe Veira Roque

Um Framework Conceitual para análise e comparação de padrões em apresentações orais de estudantes

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Roberto Muñoz, Dr.
Universidade de Valparaíso

Prof. Tiago Primo, Dr.
Universidade Federal de Pelotas

Prof. Vinicios Faria Culmant Ramos, Dr.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Tecnologias da Informação e Comunicação.

Coordenação do Programa de
Pós-Graduação

Prof. Cristian Cechinel, Dr.
Orientador

Araranguá, 2020.

Este trabalho é dedicado à minha esposa, Jaquelini, que durante todo o processo do trabalho esteve presente dando o suporte necessário para a conclusão do mesmo.

AGRADECIMENTOS

À minha família, que durante todo o período do curso deu todo o apoio necessário para que o mesmo fosse possível. À Universidade Federal de Santa Catarina, pelas oportunidades que me foram concedidas. Ao professor e orientador Cristian Cechinel, que durante todo esse período foi um guia para novos conhecimentos e um profissional dedicado. Ao professor Roberto Muñoz que forneceu grande parte do material da pesquisa e auxiliou em grande parte do estudo. À todos os professores da universidade pelos conhecimentos me passados durante o curso. Aos meus amigos que durante toda a realização desse trabalho me deram apoio e ajuda em diversos aspectos do trabalho. À minha esposa, Jaqueline, pelo apoio, motivação e compreensão nesse período de execução e finalização do trabalho. Aos demais colegas e amigos que estiveram presentes nesse aprendizado constante ao longo dos anos na universidade.

*"If I have seen further
it is by standing on
the shoulders of Giants."
(Isaac Newton)*

RESUMO

Falar e apresentar em público é uma habilidade essencial para o desenvolvimento acadêmico e profissional. Essas habilidades são exigidas em toda a sociedade e seu desenvolvimento e avaliação são um desafio enfrentado pelas instituições de ensino superior. Existem alguns desafios para avaliar objetivamente, além de gerar informações valiosas para os professores e feedback apropriado para os alunos. Neste estudo, um Framework conceitual é apresentado para entender e detectar padrões nas apresentações orais dos alunos. São propostas 4 etapas para o Framework: Coleta de dados, Análise Estatística, *Clustering*, e a etapa de *Sequential Pattern Mining*. Na Coleta de Dados as interações durante as apresentações são coletadas e arranjadas de forma a permitir análises posteriores. O bloco de Análise Estatística possibilita uma compreensão geral dos dados coletados, pode evidenciar diferenças ou semelhanças entre apresentações ou anos, dependendo dos dados coletados. A etapa de *Clustering* permite a segmentação dos estudantes em grupos com características bem definidas. Por fim, a análise por *Sequential Pattern Mining* pode atuar como complementadora dos resultados anteriores e ainda permitir identificar sequências que caracterizem as apresentações. Um estudo de caso foi realizado, foram coletados dados de 222 estudantes entrantes de Engenharia da Computação (CE) em três momentos diferentes, durante dois anos diferentes. As análises possibilitaram segmentar os apresentadores em grupos distintos com suas características diferenciando seus comportamentos. Testes inferenciais permitiram avaliar a evolução do apresentador ao longo de cada ano, apontando uma convergência em termos de redução do número de atributos estatisticamente diferentes entre as apresentações realizadas no mesmo momento do curso. A análise sequencial permite observar as sequências posturais mais frequentes, possibilitando ainda uma outra perspectiva de avaliação dos dados. Os resultados podem ajudar a retornar aos alunos um *feedback* automático em termos de suas posturas e discursos ao longo das apresentações, podendo ainda servir como informações de base para futuras comparações com apresentações de estudantes provenientes de diferentes cursos de graduação.

Palavras-chave: *Multimodal Learning Analytics*. Apresentações Orais. Mineração de Dados Educacionais.

ABSTRACT

Public speaking and presenting is an essential skill for academic and professional development. These skills are required across society and their development and evaluation is a challenge faced by higher education institutions. There are some challenges to assess objectively, in addition to generating valuable information for teachers and appropriate feedback for students. In this study, a conceptual framework is presented to understand and detect patterns in students' oral presentations. Four steps are proposed for the Framework: Data collection, Statistical Analysis, *Clustering*, and the *Sequential Pattern Mining* step. In Data Collection, interactions during presentations are collected and arranged in order to allow for further analysis. The Statistical Analysis block provides a general understanding of the data collected, it can show differences or similarities depending on the data collected. The stage of *Clustering* allows the segmentation of students into groups with well-defined characteristics. Finally, the analysis by *Sequential Pattern Mining* can complement the previous results and also allow the identification of sequences that characterize the presentations. A case study was carried out, data were collected from 222 incoming students of Computer Engineering (CE) at three different times, during two different years. The analyzes made it possible to segment the presenters into distinct groups with their characteristics, differentiating their behaviors. Inferential tests allowed to assess the presenter's evolution throughout each year, pointing to a convergence in terms of reducing the number of statistically different attributes between the presentations made at the same moment in the course. Sequential analysis makes it possible to observe the most frequent postural sequences, enabling yet another perspective for data evaluation. The results can help to give students an automatic *feedback* in terms of their postures and speeches throughout the presentations, and can also serve as background information for future comparisons with presentations from students from different undergraduate courses.

Keywords: Multimodal Learning Analytics. Oral Presentations. Educational Data Mining.

LISTA DE FIGURAS

Figura 1 – Usando k-means nas posturas corporais de estudantes. Primeiro estado (esquerda) ativo, duas mãos na mesa; o segundo (meio) passivo, com os braços cruzados; terceiro (direita) semi-ativo, somente uma mão sobre a mesa. Fonte:(SCHNEIDER; BLIKSTEIN, 2015)	25
Figura 2 – Posturas agrupadas do tronco de estudantes apresentando. Fonte: (ECHEVERRIA <i>et al.</i> , 2014)	26
Figura 3 – Hierarquia de técnicas de aprendizado de máquina. Fonte: Elaborado pelo autor.	28
Figura 4 – Framework Conceitual proposto. Fonte: Elaborado pelo autor.	34
Figura 5 – Configuração do ambiente. Fonte:(ROQUE, F. <i>et al.</i> , 2019b)	39
Figura 6 – Tela do aplicativo Leikelen mostrando o modelo de <i>skeleton</i> de duas pessoas.	41
Figura 7 – Gráficos de tendência comparando os anos.Fonte: Adaptado de (ROQUE, F. <i>et al.</i> , 2019b).	46
Figura 8 – Gráficos das <i>Silhouettes</i> para três diferentes momentos em 2017.Fonte: Adaptado de (ROQUE, F. <i>et al.</i> , 2019b)	51
Figura 9 – Gráficos das <i>Silhouettes</i> para três diferentes momentos em 2017.Fonte: Adaptado de (ROQUE, F. <i>et al.</i> , 2019b)	52
Figura 10 – Gráficos polares dos centroides de cada apresentação. Fonte: (ROQUE, F. <i>et al.</i> , 2019b)	55

LISTA DE TABELAS

Tabela 1 – Base de dados sequencial.	32
Tabela 2 – Detalhes dos processos do Framework agrupados. Fonte: Elaborado pelo autor.	36
Tabela 3 – Descrição de cada atributo das bases de dados.	42
Tabela 4 – Número de observações de cada base de dados.	42
Tabela 5 – Valor médio para cada atributo.	43
Tabela 6 – 5 atributos mais presentes de cada base de dados.	44
Tabela 7 – Testes inferenciais entre as bases de dados. Fonte: (ROQUE, F. <i>et al.</i> , 2019b)	48
Tabela 8 – Média dos valores das <i>Silhouettes</i>	50
Tabela 9 – Sequências selecionadas da primeira apresentação de 2017.	58
Tabela 10 – Sequências selecionadas da segunda apresentação de 2017.	59
Tabela 11 – Sequências selecionadas da terceira apresentação de 2017.	59
Tabela 12 – Sequências selecionadas da primeira apresentação de 2018.	60
Tabela 13 – Sequências selecionadas da segunda apresentação de 2018.	60
Tabela 14 – Sequências selecionadas da terceira apresentação de 2018.	61

LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina
DS	Data Storytelling
IA	Inteligência Artificial
JVA	Joint Visual Attention
LA	Learning Analytics
LMS	Learning Management System
MMLA	Multimodal Learning Analytics
NISPI	Nonverbal indexes of students' physical interactivity
SDK	Software Development Kit
SPMF	quential Pattern Mining Framework
SVM	Supported Vector Machine
TUI	Tangible User Interface

SUMÁRIO

1	INTRODUÇÃO	14
1.1	CONTEXTUALIZAÇÃO DO PROBLEMA E JUSTIFICATIVA DA PESQUISA	14
1.2	OBJETIVOS	16
1.2.1	Objetivo Geral	16
1.2.2	Objetivos Específicos	16
1.3	ADERÊNCIA DO OBJETO DE PESQUISA AO PPGTIC	16
1.4	ORGANIZAÇÃO DO TRABALHO	17
2	ESTADO DA ARTE	18
2.1	<i>MULTIMODAL LEARNING ANALYTICS</i> NO CONTEXTO DA SALA DE AULA	18
2.2	<i>MULTIMODAL LEARNING ANALYTICS</i> PARA <i>FEEDBACK</i> EM APRESENTAÇÕES ORAIS	19
2.3	<i>MULTIMODAL LEARNING ANALYTICS</i> , MODELOS E FRAMEWORKS	20
3	FUNDAMENTAÇÃO TEÓRICA	21
3.1	FRAMEWORK CONCEITUAL	21
3.2	<i>MULTIMODAL LEARNING ANALYTICS</i>	21
3.2.1	Modalidades de Dados	23
3.2.1.1	Olhar	23
3.2.1.2	Posturas e Gestos	24
3.2.1.3	Análise de Texto e Desenhos	25
3.2.1.4	Fala	26
3.3	APRENDIZADO DE MÁQUINA E ESTATÍSTICA	27
3.3.1	Aprendizados Supervisionado e não Supervisionado	28
3.3.1.1	<i>K-means</i>	29
3.3.1.2	<i>Silhouettes</i>	30
3.3.2	Estatística Inferencial	30
3.4	<i>SEQUENTIAL PATTERN MINING</i>	31
4	PROCEDIMENTOS METODOLÓGICOS	34
4.1	FRAMEWORK CONCEITUAL PARA AVALIAÇÃO DE PADRÕES EM POSTURAS CORPORAIS DE APRESENTAÇÕES ORAIS	34
4.1.1	Coleta de Dados - A	35
4.1.2	Análise Estatística - B	37
4.1.3	Clustering - C	37
4.1.4	<i>Sequential Pattern Mining</i> - D	38
4.2	ESTUDO DE CASO	38
4.2.1	Ambiente de Aprendizagem das Apresentações	38

4.2.1.1	Descrição do Caso	39
4.2.1.2	Microsoft Kinect	39
4.2.1.3	Procedimentos e Coleta de Dados	40
4.2.1.4	Ferramenta Leikelen	40
4.2.1.5	Características Coletadas (Atributos)	41
4.2.1.6	Descrição dos Dados	42
5	RESULTADOS - ESTUDO DE CASO	45
5.1	COMPARAÇÃO DOS ATRIBUTOS ENTRE OS CURSOS	45
5.2	GRUPOS (CLUSTERIZAÇÃO)	48
5.2.1	<i>Os Grupos (Clusters)</i>	49
5.3	COMPARAÇÃO ENTRE OS CLUSTERS - ANÁLISE DOS CENTROI- DES	50
5.3.1	Apresentações de 2017	53
5.3.2	Apresentações de 2018	54
5.3.3	Comparação Entre os Anos	55
5.4	<i>SEQUENTIAL PATTERN MINING (SPM)</i>	57
5.4.1	Aplicação do MINSUP 95% nas Apresentações de 2017	57
5.4.2	Aplicação do MINSUP 60% nas Apresentações de 2018	59
5.5	DISCUSSÃO	61
6	CONCLUSÃO	64
6.1	TRABALHOS FUTUROS	65
	REFERÊNCIAS	67
	APÊNDICE A – PUBLICAÇÕES	79
A.1	PUBLICAÇÕES	79

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO DO PROBLEMA E JUSTIFICATIVA DA PESQUISA

O desenvolvimento das habilidades de comunicação e trabalho em equipe tem sido indicado por organizações profissionais e recomendações de currículos e relatórios de programas de graduação universitária. Estas instituições indicam essas habilidades como conhecimento fundamental para a preparação de futuros profissionais em diferentes áreas do conhecimento. Em relação às habilidades de comunicação, o objetivo principal é transmitir as informações de forma clara e coerente para as pessoas (LUCAS, 2011). Bons comunicadores demonstram suas habilidades de apresentação com características verbais e não verbais, como linguagem corporal, contato visual com o público ou o espaço que eles ocupam no palco (YORK, 2013). A linguagem corporal é uma parte importante dos processos de aprendizado e comunicação (MUNOZ *et al.*, 2018). Além disso, posturas e gestos corporais são usados em apresentações orais para transmitir ideias e mensagens (MEHRABIAN, 2017).

Enquanto que a comunicação verbal pode ser a chave em interações sociais, muitas informações são passadas por comportamento não verbal. O comportamento não verbal, como expressões faciais e movimento do corpo, comunicam mais informação sobre o que uma pessoa está pensando ou sentindo do que as palavras (RICKHEIT; STROHNER, 2008). Mesmo estimativas conservadoras apontam que o papel do comportamento não verbal é de mais de 60% (BURGOON; HOOBLER, 1994). A comunicação que ocorre somente pela forma de palavras/fala não é suficiente para transmitir a informação que se quer aos ouvintes (HYMES, 1972).

As habilidades interpessoais, como falar em público, são muito importantes para uma variedade de profissões (RASIPURAM; JAYAGOPI, 2020). A comunicação em ambientes sociais e públicos podem influenciar o desenvolvimento da carreira, construir relacionamentos, resolver conflitos, ou até mesmo ajudar em negociações (WÖRTWEIN *et al.*, 2015).

Uma apresentação de sucesso em público está associada a diversos fatores. Algumas modalidades desses fatores podem ser o conteúdo do discurso, a voz e sua entonação, as expressões faciais, posição da cabeça, gestos das mãos e posturas corporais (CHEN *et al.*, 2014). Contudo, falar ou apresentar em público não é uma habilidade nata de todos, mas pode ser dominada com treino intenso (HART *et al.*, 2013).

Extrair e processar de forma automática com áudios e vídeos características não verbais e verbais pode ser importante no processo de aprendizagem. Ainda, determinar a utilidade dessas características no problema pode indicar um caminho para entender e caracterizar o aprendiz ou envolvidos no processo (BEYAN *et al.*, 2016).

Atualmente, tecnologias e sensores avançados estão disponíveis em ambien-

tes de aprendizado que permitem capturar dados através de diferentes modalidades (NOEL *et al.*, 2018). De acordo com Ochoa (2017), o campo de Multimodal Learning Analytics (MMLA) incorpora diferentes fontes de traços de aprendizado na pesquisa de Learning Analytics (LA), além do arquivo de log tradicional, capturado pelos sistemas on-line. Tecnologias avançadas de sensores permitem capturar dados biométricos com diferentes modalidades. Isso inclui artefatos, como olhar (SCHNEIDER; PEA, 2013), posturas (OCHOA *et al.*, 2018; MUNOZ *et al.*, 2018), ações (WORSLEY; BLIKSTEIN, 2014), expressões faciais (D'MELLO *et al.*, 2008), discurso (THOMPSON, 2013; NOEL *et al.*, 2018), escrita (OVIATT; COHEN, A., 2013) e esboços/desenhos (WORSLEY; BLIKSTEIN, 2011). Esses artefatos são exemplos de informações multimodais que foram gradualmente incorporadas aos sistemas de aprendizado.

Existem alguns desafios associados à combinação de dados de várias fontes, resultando em dados ruidosos (OCHOA; WORSLEY, 2016). Problemas com dados de várias fontes e programas, envolvendo reconhecimento facial e de fala, geralmente têm resoluções complexas, que usualmente são resolvidas por meio de técnicas de Aprendizado de Máquina (AM) (OCHOA *et al.*, 2016; DUDA *et al.*, 1995). Além disso, a identificação e caracterização de estudantes de graduação é um ponto de partida fundamental, a fim de obter melhores processos de ensino e aprendizagem (RIQUELME *et al.*, 2019). Na avaliação de alunos que usam técnicas de MMLA, após a captura de dados de um grupo de usuários e fontes diferentes, os conjuntos de dados devem ser mesclados para processamento estatístico ou mineração de dados de modo a extrair informações razoáveis. As técnicas de aprendizado de máquina, que criam modelos a partir de dados passados, são importantes em áreas onde existem grandes análises de dados. O MMLA, junto com as técnicas de aprendizado de máquina, é uma abordagem poderosa que ajuda no desenvolvimento da educação - (BLIKSTEIN, 2013) e, principalmente, na análise do comportamento do aluno durante apresentações orais.

Um dos componentes centrais do MMLA contemporâneo é o desenvolvimento de Frameworks e modelos que oferecem melhor generalização e aplicabilidade (WORSLEY, 2018). Da mesma maneira, utilizar Frameworks pode ajudar a estabelecer normas em como dados são analisados em diferentes contextos. Ainda para Worsley (2018) essas ferramentas possibilitam que pesquisadores situem mais claramente os objetivos e orientação de seus trabalhos.

Este trabalho apresenta uma abordagem para a identificação de padrões na apresentação de alunos usando dados multimodais das informações de fala e corpo que foram coletadas dos alunos através do *Microsoft Kinect*. A abordagem consiste na aplicação e avaliação de técnicas de agrupamento (*Silhouettes* e *K-means*), a fim de encontrar as categorias e perfis distintos desses apresentadores. Para isso, foram considerados grupos de estudantes de graduação em um estudo de caso com dois anos diferentes (2017 e 2018), com base nos estudos experimentais de apresentações orais

durante o curso do projeto do ano acadêmico. O objetivo final é o desenvolvimento de um sistema capaz de reconhecer automaticamente os perfis dos alunos participantes durante uma apresentação oral, a fim de permitir *feedback* imediato sobre as apresentações realizadas. Para o presente estudo, as seguintes perguntas de pesquisa são propostas:

- P1: Quantos e quais são os diferentes grupos de padrões encontrados nas apresentações orais dos alunos?
- P2: Quais são as semelhanças e diferenças entre as apresentações dos diferentes anos?
- P3: É possível observar padrões nos resultados de aprendizagem de estudantes de ambos os anos em termos de posturas corporais?

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Propor um Framework Conceitual para identificação de padrões corporais e de fala de estudantes durante apresentações orais.

1.2.2 Objetivos Específicos

- Analisar artigos na literatura sobre *Multimodal Learning Analytics* no ambiente de sala de aula.
- analisar um estudo de caso com dados coletados pela Universidade de Valparaíso, Chile.
- Testar diferentes métodos de análise de dados, como: Análise de estatística descritiva e inferencial, Aprendizado de Máquina, e Mineração de Padrões Sequenciais .
- Interpretar os dados gerados no contexto educacional e propor um Framework Conceitual para análise de padrões em apresentações orais de estudantes.

1.3 ADERÊNCIA DO OBJETO DE PESQUISA AO PPGTIC

Este trabalho se relaciona diretamente com o objeto de formação do PPGTIC. O trabalho em questão abordará tanto a temática computacional quanto a educacional. Primeiramente, com um foco maior, são analisados dados por abordagens de aprendizado de máquina, se enquadrando na linha computacional. Em outra linha, dados educacionais voltados para o entendimento de apresentações orais serão usados.

Dessa forma, o trabalho está totalmente alinhado com um tema atual de *Learning Analytics* e sua sub área *Multimodal Learning Analytics*.

1.4 ORGANIZAÇÃO DO TRABALHO

O restante do trabalho está estruturado da seguinte maneira. O Capítulo 2 apresenta o estado da arte do problema. A descrição da fundamentação teórica sobre Aprendizado de Máquina e estatística, *Multimodal Learning Analytics* e Framework Conceitual é realizada no capítulo 3. Os procedimentos metodológicos adotados para a realização do trabalho são descritos/apresentados no Capítulo 4. Os resultados e discussão dos dados e informações encontradas são apresentados no Capítulo 5. Por fim, as conclusões e oportunidades para estudos futuros são apresentadas no Capítulo 6.

Ainda, os capítulos 2,4 e 5 são baseados em parte ou totalmente em materiais que já foram publicados: Felipe Roque *et al.* (2019b), Felipe Vieira Roque *et al.* (2018), Felipe Roque *et al.* (2018) e Felipe Roque *et al.* (2019a).

2 ESTADO DA ARTE

Este capítulo apresenta uma breve revisão bibliográfica acerca do tema estudado para o entendimento completo do trabalho. São apresentados estudos envolvendo o MMLA no contexto da sala de aula, estudos de MMLA para *feedback* em apresentações orais e a criação de Frameworks e modelos no campo do MMLA.

2.1 MULTIMODAL LEARNING ANALYTICS NO CONTEXTO DA SALA DE AULA

Multimodal Learning Analytics é um campo relacionado a LA com ênfase na análise de modalidades complexas de comunicação durante atividades de aprendizagem, como fala, escrita e interação não-verbal (SCHERER *et al.*, 2012). Embora relativamente jovem, o MMLA é um campo ativo e uma variedade de técnicas e aplicações foram propostas nos últimos anos (WORSLEY, 2017).

Uma abordagem usual no MMLA é a integração de gravações de câmeras com outros dados para analisar o comportamento do aluno. Por exemplo, Bidwell e Fuchs (2011) usaram uma gravação de vídeo com várias câmeras e classificações de observadores para propor uma estrutura de análise de comportamento. Seu objetivo era modelar e classificar o envolvimento dos alunos a partir de sequências de foco de olhar do aluno, usando o software de rastreamento de rosto. Além disso, um grupo de observadores especialistas classificou os comportamentos dos alunos em categorias (engajados, atentos ou em transição). Com essas duas entradas, um classificador de envolvimento do aluno foi criado para fornecer um relatório ao professor. Além disso, o estudo apresentado em (OCHOA *et al.*, 2013), mostra que informações em vídeo, áudio e toque de caneta foram usadas para discriminar entre especialistas e não especialistas em grupos de estudantes que resolvem problemas matemáticos.

Outra instância de gravações de vídeo com várias câmeras sendo usadas em contextos de sala de aula é o trabalho de Raca e Dillenbourg (2013), em que 3 a 4 câmeras foram usadas para coletar dados em uma sala de aula comum, de maneira discreta. Em seguida, as informações sobre o olhar e o movimento corporal foram extraídas dos vídeos para fornecer um relatório ao professor. Em (RACA *et al.*, 2014), além das imagens coletadas, questionários, entrevistas e um rastreador ocular também foram utilizados no experimento.

No contexto mais específico dos padrões posturais, muitos trabalhos anteriores usaram câmeras tradicionais para coletar dados - (COHEN, I.; LI, 2003; MO *et al.*, 2009; ZHAO *et al.*, 2011; CHELLA *et al.*, 2005; MOGHADDAM; PICCARDI, 2010). Usando um dispositivo Kinect, no estudo de Le, Nguyen *et al.* (2013) foi proposto uma alternativa para resolver problemas que de outra forma seriam difíceis de resolver usando câmeras convencionais no reconhecimento de posturas humanas. Eles usaram a ferramenta de rastreamento esquelético fornecida pelo Kinect Software Development Kit

(SDK) para coletar informações conjuntas. Com base no esqueleto rastreado, posições e ângulos articulares foram utilizados como características. Estes foram então utilizados na técnica de Supported Vector Machine (SVM) para reconhecer posturas humanas. Munoz *et al.* (2018) desenvolveram uma ferramenta para capturar, classificar (usando o algoritmo AdaBoost) e visualizar as posturas corporais dos alunos, usando um dispositivo Kinect, permitindo também o uso de dados de outros sensores.

Echeverria *et al.* (2014) também usaram o dispositivo Microsoft Kinect para detectar posturas comuns das articulações do corpo do esqueleto. Eles então empregaram o *Fuzzy C-Means* para extrair generalizações de posturas, por pessoa e, em seguida, realizaram clusterização usando o algoritmo *K-Means*. Usando um simulador que permitia os estudantes resolverem problemas da Lei de Ohm, o trabalho Zhu *et al.* (2019) apresenta uma abordagem de *Sequential Pattern Mining* para entender a relação entre os estudantes que desenvolveram o problema com sucesso e os que não obtiveram êxito. O trabalho também apresenta uma análise de taxa de transições e de entropia para chegar nas conclusões do estudo.

Reilly *et al.* (2018) analisaram o Kinect e dados de fala, coletados durante uma tarefa de programação de robôs (REILLY *et al.*, 2018). Eles mostraram a correlação entre certos movimentos e padrões de gestos com ganhos de aprendizado. Eles também usaram algoritmos de agrupamento para encontrar categorias de posições corporais recorrentes, que foram usadas para analisar a quantidade de tempo que os alunos gastaram com eles e os ganhos de aprendizado.

2.2 MULTIMODAL LEARNING ANALYTICS PARA FEEDBACK EM APRESENTAÇÕES ORAIS

O MMLA também foi usado para fornecer *feedback* de apresentação oral para professores e alunos. As apresentações são uma maneira importante para os seres humanos transmitirem informações a grandes grupos. Como consequência, os alunos são frequentemente avaliados em relação a essas apresentações. Uma abordagem ao *feedback* da apresentação oral é usar apenas sensores de baixo custo, para fornecer dados de áudio e vídeo (OCHOA *et al.*, 2018) ou de áudio e slides (LUZARDO *et al.*, 2014). Os dispositivos Kinect também podem ser usados para coletar dados para *feedback* da apresentação. No trabalho mencionado anteriormente de Echeverria *et al.* (2014), um conjunto de preditores baseados em dados de vídeo e Kinect foram usados para estimar as habilidades de apresentação dos alunos. A partir dos registros de vídeo, eles estimaram o olhar do apresentador e outros valores descritivos. A partir dos registros do Kinect, os autores extraíram posturas e características comuns das articulações corporais das informações do esqueleto.

O trabalho de Gan *et al.* (2015), um uso mais intensivo de sensores exigia câmeras tradicionais, registros de Kinect e sensores *Google Glass* para criar uma

estrutura de análise de apresentação com vários sensores. Outro trabalho extenso no uso no MMLA pode ser encontrado em Chan *et al.* (2020). Neste trabalho foram realizadas gravações de aulas práticas de alunos com câmeras em diferentes ângulos. No estudo, os autores reforçam que a informação da coleta e processamento dos dados é usada para explicar e prever níveis altos de construções, como atenção e engajamento dos estudantes.

2.3 MULTIMODAL LEARNING ANALYTICS, MODELOS E FRAMEWORKS

Para Spikol *et al.* (2018), um dos grandes desafios do MMLA atual pode ser criar Frameworks flexíveis de modo que seja possível integrar essas ferramentas em Learning Management System (LMS). Esses Frameworks ou modelos podem indicar maneiras tanto na coleta quanto na avaliação dos dados coletados em variadas modalidades de dados.

Em Cukurova *et al.* (2018) é apresentado um Framework que usa índices não verbais da interatividade física dos alunos, em inglês, Nonverbal indexes of students' physical interactivity (NISPI) em práticas baseadas em problemas. No trabalho os autores mostram uma maneira original de identificar comportamentos diferentes em resolução colaborativa de problemas.

Echeverria *et al.* (2018) empregaram Data Storytelling (DS) para melhorar a visualização de dados e como o MMLA é comunicado aos alunos e educadores. No estudo, é proposto um modelo conceitual que aplica princípios de DS em questões colocadas em *dashboards* voltados para professores e alunos. Esse tipo de resultado pode fornecer *feedback* no contexto educacional.

Focados numa área pouco explorada, Martinez-Maldonado *et al.* (2018) abordam a aprendizagem no contexto de tarefas de aprendizagem física / motora e nos aspectos da fisicalidade das tarefas intelectuais "tradicionais" que ocorrem frequentemente nos espaços de aprendizagem física. Eles estão desenvolvendo um Framework Teórico que usa a teoria da cognição distribuída, a Internet das Coisas e a análise de aprendizagem multimodal.

Os estudos aqui apresentados apresentam uma grande gama de abordagens, usando varias técnicas e métodos. Contudo, nenhum dos estudos apresenta uma abordagem sequencial para avaliar apresentações orais de estudantes. Isso pode ser notado como uma diferença entre os estudos apresentados e o estudo proposto no trabalho. Além disso, a base de dados se mostra ampla em relação a informações e numero de observações, o que também pode ser considerado como ponto diferente em relação aos trabalhos.

3 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os aspectos teóricos de alguns conceitos necessários para o entendimento completo do trabalho. São apresentados alguns conceitos referentes a Framework Conceitual, *Multimodal Learning Analytics*, Aprendizado de Máquina e estatística, e *Sequential Pattern Mining*.

3.1 FRAMEWORK CONCEITUAL

Muitos problemas são complexos e ligados a muitas áreas do conhecimento que pertencem a diferentes disciplinas. Assim, uma melhor compreensão desses problemas requer uma abordagem multidisciplinar (JABAREEN, 2009). Métodos qualitativos e quantitativos servem como ferramenta adequada para investigar esses fenômenos complexos. Dessa forma, criar um modelo de modo a investigar um problema definido pode ajudar a esclarecer as questões de pesquisa e servir de guia para o estudo.

Um Framework Conceitual pode ser usado quando o pesquisador não consegue explicar sua questão de pesquisa inteiramente em referências prévias em apenas uma teoria, ou conceitos residentes em uma teoria (IMENDA, 2014). Nesses casos, o pesquisador deve sintetizar na literatura referente a um problema, tanto as descobertas teóricas quanto as empíricas. A síntese pode ser chamada de modelo ou Framework Conceitual, que representa essencialmente uma maneira integrada para olhar um problema (LIEHR; SMITH, 1999).

Para Jabareen (2009), o Framework Conceitual pode ser uma rede, ou um plano, de conceitos interligados que juntos fornecem uma compreensão abrangente de um fenômeno ou fenômenos. Os conceitos que constituem um Framework Conceitual suportam um ao outro, articula os fenômenos respectivos, e estabelece uma filosofia específica do Framework. Já para Rocco e Plakhotnik (2009), como um tipo de manuscrito, um Framework Conceitual relaciona conceitos, pesquisa empírica, e teorias relevantes para avançar e sistematizar o conhecimento sobre conceitos ou problemas relacionados.

3.2 MULTIMODAL LEARNING ANALYTICS

Na sua origem, o foco no campo de *Learning analytics* era o estudo de performances de estudantes usando alguma forma digital de ferramenta (OCHOA, 2017). Estas ferramentas digitais poderiam ser sistemas de gerenciamento de aprendizagem, tutores inteligentes, cursos massivos online, games educativos ou qualquer tipo de sistema que use um computador como um componente ativo no processo de aprendizagem (OCHOA, 2017).

Os sistemas de aprendizagem baseados em computação, mesmo quando não

são criados com intuito de "analisar dados" em primeira mão, capturam automaticamente em grandes detalhes as interações do usuário. Essas interações são armazenadas em muitas formas, como: arquivos de log ou arquivos de texto que pode em uma segunda etapa ser minerados.

Ainda para (OCHOA, 2017), o LA falha em ambientes onde computadores não podem ou não são utilizados. Ambientes ou situações em que as ações dos usuários não podem ser automaticamente capturadas pelo sistema computacional de aprendizagem. Por exemplo, mesmo quando um estudante está resolvendo um problema em um tutor inteligente, as expressões do estudante não são capturadas por LA tradicional, mesmo sendo informação valiosa na caracterização da aprendizagem (OCHOA, 2017).

O MMLA, sendo um subcampo do LA, atua incorporando diferentes fontes de dados de aprendizes, focando na interpretação e na otimização do aprendizado em cenários reais e digitais onde as interações não são necessariamente através de um computador ou um dispositivo digital (BLIKSTEIN, 2013). Para Worsley *et al.* (2016), o MMLA se estabelece na relação entre três áreas, são elas: o ensino e o aprendizado multimodal, dados multimodais e as análises suportadas por computadores. Segundo os autores, o MMLA em essência utiliza dados tradicionais e não tradicionais com o objetivo de caracterizar ou modelar o aprendizado do estudante em ambientes complexos de aprendizagem.

Os dados utilizados no MMLA podem ser gravações de áudio e vídeos, traços de caneta, sensores de posição, movimento e profundidade, bio sensores e qualquer modalidade que pode ser útil na compreensão e na medição do processo de aprendizagem. Combinando as diferentes modalidades de dados pode-se melhorar a compreensão do estado interno do aprendiz (OCHOA, 2017).

Worsley e Blikstein (2013) definem o MMLA como sendo um conjunto de sensores multimodais que podem ser usados para prever, entender e quantificar o aprendizado dos estudantes. Em sua essência o MMLA visa alavancar dados de modalidades não tradicionais de forma a estudar e analisar o aprendizado de estudantes em ambientes de aprendizado complexos (WORSLEY, 2017). Essas técnicas permitem que professores deem retorno aos estudantes nas relações de aprendizado. De outro lado, técnicas de inteligência artificial estão cada vez mais presentes, principalmente quando se tratam de grandes quantidades de dados. Nesse sentido, as apresentações orais são tema de estudo, onde pode-se avaliar as apresentações de estudantes. Uma vez que essas apresentações são avaliadas, é possível dar um retorno ao aluno de forma que este consiga entender quais pontos devem melhorar.

A linguagem corporal é uma peça importante nos processos de aprendizagem e comunicação. As posturas corporais e os gestos são utilizados em apresentações orais para transmitir ideias e mensagens. Dessa forma, capturar e analisar comportamentos

não-verbais, ligados às posturas e aos gestos, é um importante aspecto no *Multimodal Learning Analytics* (LEONG *et al.*, 2015). Para Leong *et al.* (2015) a comunicação entre pessoas por uma causa específica, como em uma apresentação, requer cuidado para transmitir o que se quer para a parte que está ouvindo ou olhando.

3.2.1 Modalidades de Dados

Existem diferentes abordagens para acessar o aprendizado dos estudantes, afeto e fisiologia, e acessar as intenções dos estudantes (BLIKSTEIN; WORSLEY, 2016). Na definição da teoria de comunicação, a multimodalidade refere-se no uso de diversos modos de comunicação (textual, linguística, espacial, visual entre outros) para trocar informação e sentido entre indivíduos (KRESS; VAN LEEUWEN, 2001). O ser humano pode aprender lendo um livro, escutando um professor, assistindo um procedimento, usando meios digitais ou físicos, e qualquer outro meio de comunicação humano onde informação relativamente complexa pode ser gravada (OCHOA, 2017). Para Jewitt (2006), por sua própria natureza, o aprendizado é sempre multimodal.

Esta seção descreve e apresenta algumas modalidades de dados do *Multimodal Learning Analytics*. Para cada modalidade, uma breve explicação sobre o tema é apresentada. A lista de modalidades a seguir contém algumas das abordagens mais comuns de captura e gravação, e que de alguma forma contribuem para o entendimento do presente trabalho. Ressalta-se que a lista não engloba todas as modalidades presentes no MMLA.

3.2.1.1 Olhar

Humanos tendem a olhar diretamente para o que atrai sua atenção. A direção do olhar de um indivíduo é um indicativo substituto de onde está a atenção dele (FRISCHEN *et al.*, 2007). A captura do olhar tem sido utilizada dentro do campo de pesquisa de consumo de eletrônicos e uso de softwares (BLIKSTEIN; WORSLEY, 2016). Contudo, mais recentemente, em variados ambientes de aprendizado tem sido mostrado que a captura do olhar pode ser útil no entendimento do aprendizado de estudantes. Gomes *et al.* (2013) capturaram dados de olhar de estudantes de ensino médio enquanto eles completavam uma coleção de jogos de design de engenharia. Usando Aprendizado de Máquina, os pesquisadores agruparam os padrões dos olhares (baseados em atenção), identificando que as melhores performances tiveram padrões similares de foco do olhar.

Em outro trabalho, Mason *et al.* (2013) sugerem que dados de olhar também podem ajudar a entender qual tipo de abordagens são úteis para ajudar estudantes a melhorar sua aprendizagem. Os autores mostram que usar figuras em textos científicos pode ser melhor do que usar somente texto.

Entretanto, De Koning *et al.* (2010) argumentam que olhar para específicos estímulos pode representar a mudança de atenção do aluno para possíveis áreas de interesse, mas nem sempre significa que ele está aprendendo. No estudo, estudantes olharam mais e mais distante frequentemente para uma animação de instrução com pistas do que uma sem pistas. Os autores não concluem que deixar pistas reduz a carga cognitiva dos estudantes ou mesmo aumenta o aprendizado conceitual (DE KONING *et al.*, 2010).

Um promissor uso de captura de olhar na educação tem sido estudar pequenos grupos de estudo colaborativos. O modelo para este tipo de trabalho seria sincronizar dois capturadores de olhar e registrar o número de vezes que um grupo particular alcança a atenção visual comum - Joint Visual Attention (JVA) (BLIKSTEIN; WORSLEY, 2016). Em colaboração remota, Jermann *et al.* (2011), descrevem como bons programadores tendem a ter uma elevada recorrência de JVA quando estão em interações produtivas, comparados com programadores com menos habilidades. Schneider e Blikstein (2015) mostram que usando capturadores de olhar móveis e algoritmos de visão computacional, o JVA mostrou-se um importante preditor de ganhos de aprendizagem e performance em uma tarefa de resolução de problema.

3.2.1.2 Posturas e Gestos

O uso de reconhecimento de gestos tem sido amplamente utilizado como ferramenta no *Multimodal Learning Analytics*. Na comunidade de Visão Computacional alguns trabalhos nesse sentido têm sido desenvolvidos. Em Weinland *et al.* (2006), os autores demonstraram a habilidade para detectar ações básicas de pessoas relacionadas com movimento. Eles desenvolveram uma técnica que capturava as ações dos usuários independente do tamanho do corpo e do ponto de vista.

Alguns estudos usam câmeras de infra vermelho, que podem ser uma boa opção para evitar problemas relacionados com a câmera, como posição, luz e outras formas de variação visual. Schlömer *et al.* (2008) usam reconhecimento de gestos capturando e processando dados de acelerômetro de um *Nintendo Wiimote*.

Trabalhos mais recentes fazem uso de Microsoft Kinect e de sensores de infra vermelho como ferramentas de baixo custo para capturar e estudar gestos humanos. Howison *et al.* (2011) desenvolveram o *Mathematical Imagery Trainer*, que usa gestos das mãos capturados pelo Kinect como uma forma de estudar o entendimento dos estudantes sobre proporções. Os estudantes usam as mãos para indicar o relacionamento entre dois valores. O sistema também permite que os professores forneçam *feedback* imediato e troquem o tipo de instrução a medida que eles percebem as dificuldades do estudante.

O trabalho de Schneider e Blikstein (2015) usa um kinect para avaliar as estratégias dos estudantes quando estão interagindo com um Tangible User Interface (TUI).

A tarefa deles era aprender sobre o sistema auditivo humano interagindo com órgãos impressos em 3D. Com Algoritmos de Clusterização, os autores descobriram que as posturas dos estudantes caíram em três protótipos de posições: ativo, semi-ativo e passivo, como pode ser visto na Figura 1. Para os autores, a quantidade de tempo gasto no estado ativo esteve significativamente correlacionado com altos ganhos de aprendizado e quantidade de tempo gasto no estado passivo está significativamente correlacionado com baixos ganhos de aprendizado.

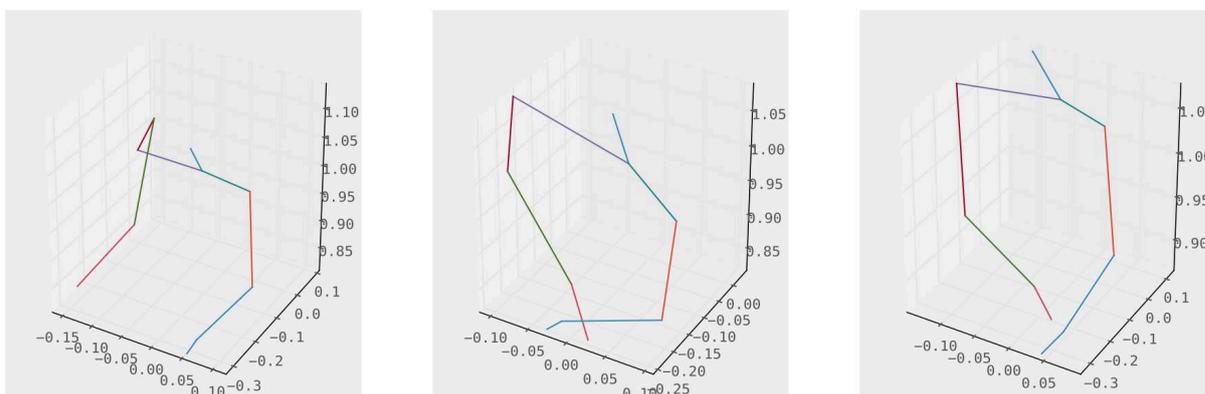


Figura 1 – Usando k-means nas posturas corporais de estudantes. Primeiro estado (esquerda) ativo, duas mãos na mesa; o segundo (meio) passivo, com os braços cruzados; terceiro (direita) semi-ativo, somente uma mão sobre a mesa. Fonte:(SCHNEIDER; BLIKSTEIN, 2015)

Em Echeverria *et al.* (2014) apresenta-se um trabalho interessado em entender e processar as posturas corporais. Analisando os gestos de estudantes novatos em apresentações orais por meio da criação de um tutor automático de apresentações. A Figura 2 mostra as posturas encontradas. No trabalho foi usado a análise dos dados do Kinect das apresentações dos estudantes. 23 posturas foram classificadas em 6 gestos corporais, que podem ser consideradas ruins ou boas para uma apresentação.

3.2.1.3 Análise de Texto e Desenhos

A análise de texto tem sido amplamente utilizada por muitas décadas em outras áreas. Mais recentemente, ela tem sido usada na educação com os pesquisadores usando texto explicitamente (BLIKSTEIN; WORSLEY, 2016). Ainda para Blikstein e Worsley (2016), o texto por si só não pode ser considerado como multimodal, a análise de texto permite a interpretação de tarefas de escrita abertas, diferentemente dos testes de múltipla escolha. A coleta de texto pode ser considerada fácil, constituindo uma das áreas mais promissoras do MMLA.

Sherin (2013), apresenta um estudo na análise de texto na área de aprendizado. O trabalho usa técnicas de modelagem e clusterização para explicar a progressão dos estudantes enquanto eles descrevem a existência das 4 estações. Ao passo que os estudantes explanam sobre a existência das 4 estações usando diferentes tipos de

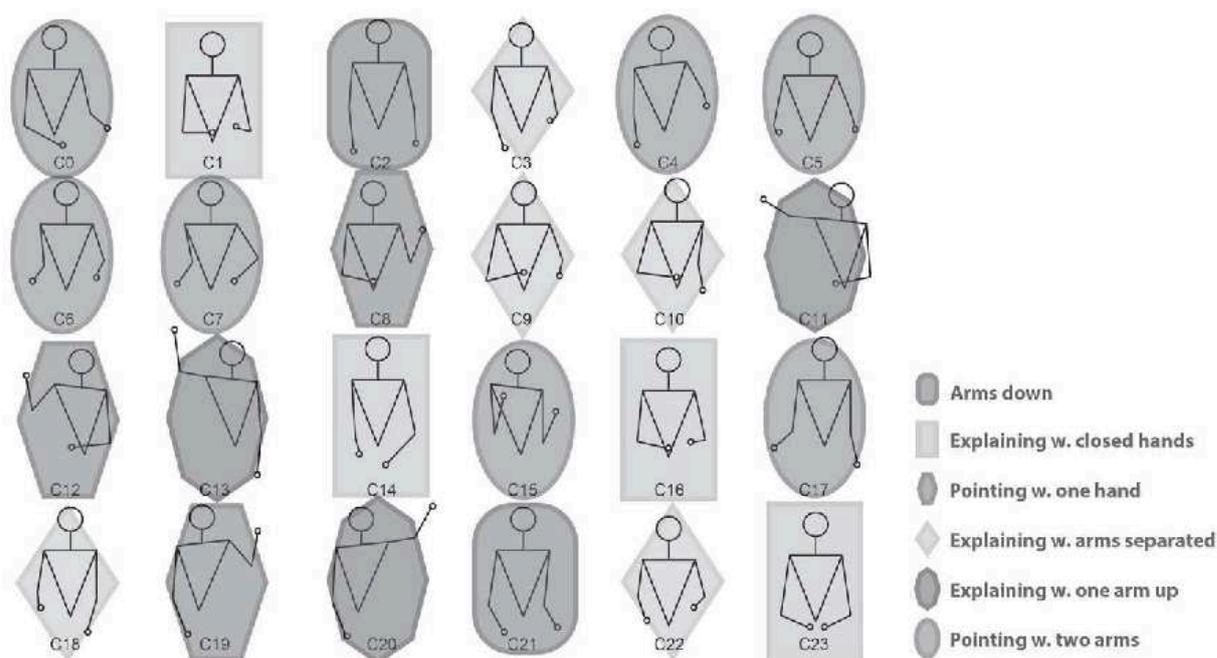


Figura 2 – Posturas agrupadas do tronco de estudantes apresentando. Fonte: (ECHEVERRIA *et al.*, 2014)

explicações científicas, é possível identificar qual tipo de explicação cada estudante se refere a cada ponto diferente do tempo (SHERIN, 2013).

Para Ochoa (2017), usar uma caneta é provavelmente uma das primeiras habilidades que estudantes aprendem, usando para escrever e desenhar. Usar meios tecnológicos para capturar o processo de desenhar e escrever possibilita "observar" informação que os sentidos humanos não são capazes de detectar facilmente, como velocidade de escrita, ritmo e pressão.

Uma maneira de registrar essas interações são as canetas digitais. Elas são capazes de digitalizar a posição, duração e pressão durante o processo de escrita. Oviatt e Adrienne Cohen (2013) exploram a contribuição que escrever e desenhar têm sobre a habilidade do indivíduo. Em um estudo similar, Zhou *et al.* (2014), usam técnicas de aprendizado de máquina supervisionado baseados em modelos de escrita e desenho para identificar especialistas em um grupo.

3.2.1.4 Fala

A forma mais comum do uso de gravações de áudio no MMLA é a captura de traços do que o estudante está falando ou escutando (OCHOA, 2017). Como a principal e mais complexa forma de comunicação entre humanos, a fala se torna importante para entender o processo de aprendizagem. Existem abordagens para avaliar o que está sendo falado, com Processador de Linguagem Natural, e também existem técnicas que avaliam como está sendo falado, como entonação, por exemplo.

Em Worsley e Blikstein (2011), os autores estudam como elementos de fala de

estudantes, como inferido por características linguísticas, textuais e prosódicas, pode identificar previamente o nível de habilidade de estudantes em tarefas de projeto de engenharia abertos.

Alguns autores conseguiram demonstrar um método para usar o reconhecimento de fala para acessar a proficiência em leitura de estudantes, como pode ser visto em Beck e Sison (2006). Completando um projeto antigo, eles combinaram reconhecimento de fala com rastreamento de conhecimento. O trabalho possui um modelo generalizado de linguagem que se tornou restrito ao conteúdo de cada livro que era estudado, o trabalho necessário para fazer reconhecimento automático de fala e subsequente classificação acurada foi bastante simplificado.

Em um estudo para melhorar o reconhecimento de fala, Munteanu *et al.* (2009) apresentam uma abordagem de palestras no âmbito universitário. Uma consideração primária na área de reconhecimento de fala, portanto, seria identificar a maneira mais efetiva de usar esta tecnologia (reconhecimento de fala) no ambiente de educação do mundo real (BLIKSTEIN; WORSLEY, 2016).

3.3 APRENDIZADO DE MÁQUINA E ESTATÍSTICA

Na computação, programas são escritos com regras bem definidas de modo a solucionar determinados problemas. Existem muitas tarefas que podem ser descritas e solucionadas por meio de um algoritmo. Contudo, existem certos tipos de problemas que têm resolução complexa por meio de programas, por exemplo, reconhecimento facial e de fala (DUDA *et al.*, 2012). No entanto, seres humanos podem resolver esse tipo de problema com muita facilidade. Os melhores reconhecedores de padrões na maioria dos casos são humanos, contudo, ainda não entendemos como os humanos reconhecem padrões (DUIN; TAX, 2005).

Na área de Inteligência Artificial (IA), técnicas de Aprendizado de Máquina têm sido utilizadas para resolver esse tipo de problema, onde é necessário reconhecer padrões em um conjunto de dados.

Quando se fala em AM, segundo Faceli *et al.* (2011) algoritmos são construídos de modo que a máquina ou o programa em si aprendam com a experiência passada. Para isso, empregam um princípio de inferência denominado indução, no qual se obtêm conclusões genéricas a partir de um conjunto genérico de dados (FACELI *et al.*, 2011). Com isso, esses algoritmos são capazes de resolver problemas a partir de dados que representam instâncias do problema a ser resolvido.

Uma definição mais formal pode ser encontrada em Russell e Norvig (2016). Um exemplo é um par $(x, f(x))$, onde x é a entrada e $f(x)$ é a saída da função aplicada a x . A tarefa da indução é: Dada uma coleção de exemplos de f , retornar uma função h que se aproxime de f .

Para Duda *et al.* (2012) existem critérios diferentes para a organização do AM.

Esses critérios dependem explicitamente da tarefa a ser realizada. São eles: Aprendizado de Máquina Supervisionado, Aprendizado de Máquina Não Supervisionado e ainda o Aprendizado por Reforço. Focando mais nos dois primeiros, enquanto um, como o nome já diz, depende de um "supervisor" externo que conhece a saída desejada. O outro explora os dados sem a necessidade de conhecimento prévio das saídas. A Figura 3 mostra melhor a hierarquia das grandes áreas do AM.

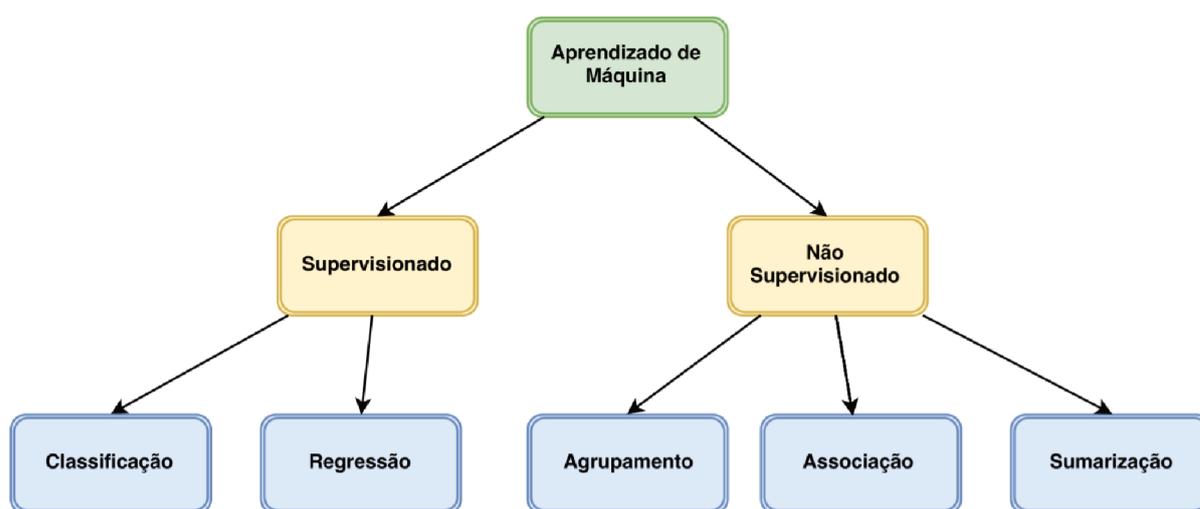


Figura 3 – Hierarquia de técnicas de aprendizado de máquina. Fonte: Elaborado pelo autor.

Mesmo que o AM seja associado a IA, outras áreas são de grande importância para o avanço do AM. A probabilidade, a estatística e a teoria da computação para citar alguns. O AM é uma das áreas que mais cresce na computação, sendo que os algoritmos usados para AM são inúmeros, com muitas variações e ainda sem restrições nas aplicações.

3.3.1 Aprendizados Supervisionado e não Supervisionado

Quando se fala em aprendizado supervisionado, uma palavra tem que estar presente, os rótulos ou etiquetas. O AM supervisionado consiste em treinar um conjunto de dados com base nesses rótulos, que podem também ser chamados de classes (DUDA *et al.*, 2012). Já o aprendizado não supervisionado não possui esses rótulos, assim, esses métodos são chamados de exploratórios. A técnica mais utilizada de AM não supervisionada é a técnica de agrupamento (DUIN; TAX, 2005).

Neste trabalho as técnicas utilizadas basicamente serão de AM não supervisionado, uma vez que os estudantes não foram categorizados previamente. O objetivo então é entender o conjunto de dados, onde grande parte do trabalho se dá por uma busca exploratória nos dados.

3.3.1.1 *K-means*

A clusterização por *K-means*, ou o algoritmo de Lloyd (LLOYD, 1982), é um algoritmo iterativo de particionamento de dados que atribui n observações para exatamente um grupo (cluster) definido por centroides. O valor de k é escolhido antes do algoritmo começar. O algoritmo funciona da maneira mostrada a seguir:

- **1.** Escolha k centros dos grupos (centroide). Podem ser escolhidos aleatoriamente.
- **2.** Calcule as distâncias do ponto ao centroide do grupo de todas as observações para cada centroide.
- **3.** Atribuir cada observação para o grupo de centroide mais próximo.
- **4.** Calcule a média das observações em cada grupo para obter k novos locais de centroide.
- **5.** Repita as etapas 2 a 4 até que as atribuições dos grupos não sejam alteradas ou o número máximo de iterações seja alcançado.

Alternativamente, o algoritmo *k-means++* usa uma heurística para encontrar sementes de centroide para o agrupamento de *k-means*. Segundo Arthur e Vassilvitskii (2006), o *k-means++* melhora o tempo de execução do algoritmo de Lloyd e a qualidade da solução final. A escolha das sementes do algoritmo *k-means++* se dá como a seguir, assumindo que o número de grupos (clusters) é k .

- **1.** Selecione uma observação aleatoriamente no conjunto de dados, X . A observação escolhida é o primeiro centroide, denotado por c_1 .
- **2.** Medir a distância de cada observação para c_1 . Denotar a distância entre c_j e a observação m dado que $d(x_m, c_j)$.
- **3.** Selecione o próximo centroide, c_2 aleatoriamente de X com a probabilidade dada pela equação 1 a seguir

$$\frac{d^2(x_m, c_j)}{\sum_{j=1}^n d^2(x_j, c_1)} \quad (1)$$

- **4.** Para escolher o centro j :
 - Medir as distâncias de cada observação para cada centroide, atribuindo cada observação para o centroide mais próximo.

- Para $m = 1, \dots, n$ e $p = 1, \dots, j - 1$, selecione o centroide j aleatoriamente de X com a probabilidade dada pela equação 2 a seguir

$$\frac{d^2(x_m, c_p)}{\sum_{\{h: x_h \in C_p\}} d^2(x_j, c_1)}, \quad (2)$$

onde C_p é o conjunto de todas as observações mais perto do centroide c_p e x_m pertencentes a C_p . Ou seja, selecionar cada centro subsequente com a probabilidade proporcional a distancia dele mesmo para o centro mais próximo que já foi escolhido.

- 5. Repetir o passo 4 até que k centroides sejam escolhidos.

3.3.1.2 Silhouettes

Para explorar os dados e analisar os grupos criados pelo *K-means*, o algoritmo de análise visual *Silhouette* (ROUSSEEUW, 1987). Esse método é útil para selecionar o melhor número de grupos (clusters) pré-definidos, por exemplo. Ele também pode ser usado para trocar uma observação que possui valor negativo da *Silhouette* para o seu vizinho, melhorando assim os resultados da análise dos grupos (ROUSSEEUW, 1987). O valor da *Silhouette* é definido pela Equação 3, como segue:

$$s(i) = \frac{b(i) - a(i)}{\max[a(i), b(i)]} \quad (3)$$

Onde $a(i)$ é a distância média entre um exemplo i e todos os outros dados dentro do mesmo grupo. Isso implica que $a(i)$ é o quanto bem determinado objeto i está dentro de seu grupo. $b(i)$ é a menor distância média de i para todos os pontos em qualquer outro grupo de quais i não faz parte. Logo, quanto mais perto de 1 o valor de $s(i)$, mais pertencente ao grupo atribuído i é. Analogamente, valores negativos representam que este exemplo não pertence ao grupo. Valores muito próximos de 0 estão no limite entre os grupo, podendo então variar entre os dois.

3.3.2 Estatística Inferencial

A estatística inferencial é o processo de análise de dados para deduzir propriedades de uma distribuição de probabilidade (UPTON; COOK, 2014). A análise estatística inferencial infere propriedades de uma população, por exemplo, testando hipóteses e derivando estimativas.

De maneira simples, os métodos de estatística inferencial, dada uma hipótese sobre uma população, para a qual deseja-se extrair inferências, a inferência estatística consiste em selecionar um modelo estatístico do processo que gera os dados e deduz proposições do modelo. Ainda, essa abordagem possibilita ao investigador argumentar das observações particulares de uma amostra para o caso geral.

Para Johnson (2011) a estatística inferencial vai muito além de descrever os dados. Ela coleta informações, interpreta e desenha conclusões. A inferência estatística inclui todos os processos de aquisição de conhecimento que envolvem a descoberta de fatos através da coleta e exame de dados. Como consequência disso, os métodos inferenciais permeiam todos os campos do conhecimento que a avaliação de informação deve ser fundamentada em evidências baseadas em dados (JOHNSON, 2011).

Nesse sentido, é importante destacar outro conceito, a estatística não paramétrica. Ela é uma ramificação da estatística não baseada somente em famílias parametrizadas de distribuições de probabilidade (parâmetros comuns são a média, variância). A estatística não paramétrica é baseada tanto em distribuições livres quanto em distribuições específicas, mas com parâmetros não específicos. Dessa forma, pela natureza do trabalho, métodos não paramétricos são necessários. Ressalta-se que antes de usar algum método os dados devem ser testados de modo a confirmar sua distribuição. O teste não paramétrico *Wilcoxon Rank Sum* (GIBBONS; CHAKRABORTI, 2011) pode ser usado para esse tipo de resultado.

3.4 SEQUENTIAL PATTERN MINING

A técnica de *Sequential Pattern Mining* é uma ferramenta conhecida na área de mineração de dados. Em uma base de dados sequencial, cada sequência consiste em uma lista de transações (SARAF *et al.*, 2015). Então, o *Sequential Pattern Mining* trabalha para encontrar todos os padrões de sequências baseado no suporte mínimo dado pelo usuário inicialmente.

Tanto dados na forma de séries temporais quanto sequências podem ser usados em problemas de mineração de padrões sequenciais. Uma sequência pode ser definida como uma lista de valores nominais ou símbolos, enquanto séries temporais são definidas como uma lista ordenada de números. As sequências e as séries podem ser usadas em muitos domínios. Para Fournier-Viger *et al.* (2017), séries temporais geralmente são usadas para representar dados como: preço de ações, temperatura e energia elétrica. Já as sequências podem representar sentenças em textos ou ainda itens comprados por um consumidor.

Uma sequência é uma lista ordenada de um conjunto de itens $s = \{I_1, I_2, \dots, I_n\}$. Por exemplo, a sequência $\langle \{a, b\}, \{c\}, \{f, g\}, \{g\}, \{f\} \rangle$, representando cinco transações feitas por um consumidor em uma loja. Cada letra sozinha representa um item. Os itens entre chaves representam um conjunto de itens. Esta sequência indica que um consumidor comprou os itens a e b ao mesmo tempo, então comprou o item c e assim por diante (FOURNIER-VIGER *et al.*, 2017).

O objetivo do SPM é encontrar sub sequências interessantes em uma base de dados sequencial, ou seja, relacionamentos entre itens que são interessantes para os usuários (FOURNIER-VIGER *et al.*, 2017). O suporte de uma sequência pode ser

Tabela 1 – Base de dados sequencial.

ID	Sequências
1	{a, b}, {c}, {f, g}, {g}, {e}
2	{a, d}, {c}, {b}, {a, b, e, f}
3	{a}, {b}, {f, g}, {e}
4	{b}, {f, g}

definido como o número de vezes que uma sequência ocorre em uma base de dados sequencial. Toma-se como exemplo o suporte da sequência $\langle\{b\}, \{f, g\}\rangle$ na Tabela 1 é 3, visto que essa sequência aparece em três sequências (1,2 e 4).

O valor do suporte também pode ser dado na forma de razão. Chamada de suporte relativo é dado pelo número de vezes que uma sequência ocorre dentro da base de dados sequencial dividido pela sequências da base de dados. A tarefa do SPM deve ser então encontrar todas as sequências frequentes em uma base de dados sequencial. Um padrão, ou sequência, é dito frequente se seu suporte é maior ou igual ao suporte mínimo dado pelo usuário.

O *quential Pattern Mining Framework (SPMF)* (FOURNIER-VIGER *et al.*, 2016) é uma biblioteca de código aberto em mineração de dados. Ela oferece mais de 120 algoritmos de mineração implementados em JAVA. A ferramenta também possui uma tela de visualização onde não é necessário trabalhar com a programação.

Inicialmente, os dados utilizados no trabalho eram sequências de *strings* onde essas continham os próprios atributos. Por exemplo, uma sequência poderia ser $\{M\tilde{a}os\ Baixas, M\tilde{a}os\ Baixas, M\tilde{a}os\ Baixas\}$, para uma determinada observação. O software SPMF não aceita esse tipo de entrada para os algoritmos, devendo essa sequência ser convertida. Convertendo o atributo *Mãos Baixas* para o número 1 (todos os atributos foram convertidos para números), a sequência que o SPMF aceita como entrada seria $\{1\ -1\ 1\ -1\ 1\ -1\ -2\}$. O -1 indica um intervalo entre um item e outro e o -2 indica o fim da sequência da observação.

Para a obtenção dos padrões sequenciais mais frequentes foi usado o algoritmo *prefixSpan*, que está presente no SPMF. Os parâmetros utilizados no algoritmo de *Minsup* variável de acordo com os resultados. O algoritmo *prefixSpan*, funciona da seguinte maneira:

- **Encontre padrões de 1-sequência** - Percorre uma base de dados sequencial *S* uma vez encontrando todos os itens frequentes em sequências, resultando no prefixos.
- **Dividir espaço de busca** - O conjunto completo de padrões sequenciais pode ser particionado baseado no prefixo derivado da primeira etapa. Todos os padrões sequenciais são derivados nessa etapa.

- **Encontre sub conjuntos de padrões sequenciais** - Os subconjuntos de padrões sequenciais podem ser extraídos através da construção do correspondente conjunto de bases de dados projetadas, sendo cada delas construída recursivamente.

4 PROCEDIMENTOS METODOLÓGICOS

Esta seção apresenta a metodologia adotada no estudo. A seção também mostra uma descrição dos dados coletados e também da base de dados utilizada. A Figura 4 apresenta um esboço geral do Framework proposto no trabalho. O bloco maior, contendo outros blocos, representa o fluxo e algumas das etapas e processos a serem realizados nos dados. O bloco menor representa os processos que devem ser feitos nos blocos maiores para obtenção dos resultados. Na sequência do framework é apresentado um estudo do caso com dados vindos da universidade de Valparaíso, Chile.

4.1 FRAMEWORK CONCEITUAL PARA AVALIAÇÃO DE PADRÕES EM POSTURAS CORPORAIS DE APRESENTAÇÕES ORAIS

Neste trabalho é proposto um Framework para avaliação de apresentações orais de estudantes ao longo de um semestre letivo. O Framework pode comparar anos diferentes e apresentações em momentos do semestre diferentes. A ferramenta usa as variáveis como forma de externalizar e evidenciar padrões que os estudantes têm durante as apresentações. Também é possível agrupar estudantes em um mesmo cluster para facilitar a avaliação dos atributos desses estudantes. As seções a seguir descrevem o funcionamento do Framework.

O Framework possui quatro grandes objetos (blocos) que juntos fornecem resultados que se conectam, encapsulando assim métodos necessários para a compreensão dos dados. De modo geral, o processo da obtenção dos resultados segue um fluxo,

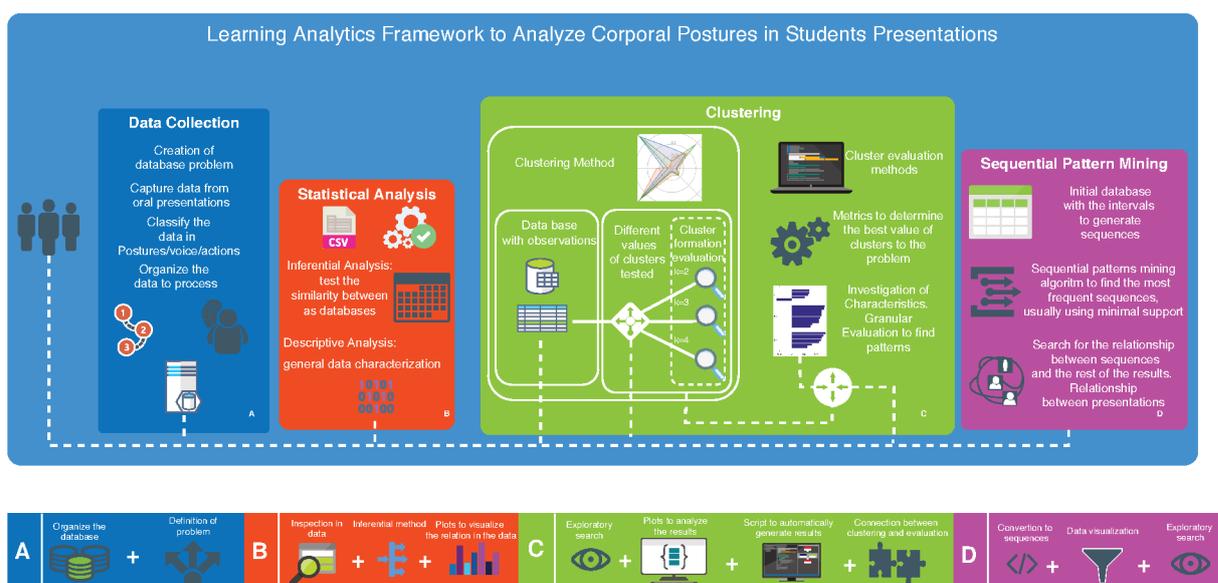


Figura 4 – Framework Conceitual proposto. Fonte: Elaborado pelo autor.

pontos brancos tracejados, contudo, em diversas etapas, como a própria montagem da base de dados é necessário retornar um passo ou mais, visto que os dados podem e muitas vezes contém ruídos que devem ser retirados, ou erros imprevistos que podem influenciar no resultado final.

O bloco A, chamado Coleta de Dados, indica como deve ser feito a coleta e pré-processamento dos dados. Essa etapa pode ser considerada a mais importante. O sucesso de todos os outros processos depende da qualidade dos dados adquiridas na coleta de dados. O bloco B, Análise Estatística apresenta meio de obter relações entre os anos e apresentações quaisquer. Com bases de dados diferentes em mãos é possível gerar relações entre elas por meio de métodos estatísticos. O bloco C, Clustering, tem por objetivo a segmentação das observações em cluster com características bem definidas. Esses resultados devem indicar comportamentos que são difíceis de visualizar sem a segmentação das observações. Ainda no mesmo intuito de evidenciar comportamentos dos estudantes, o bloco D, de Sequential Pattern Mining, deve apresentar padrões de sequências interessantes, ajudando os educadores a caracterizar as apresentações.

Cada bloco, individualmente, representa os passos, ou processos, que os dados passaram. Os blocos, por consequência, possuem seus conjuntos de resultados individuais. Por ser uma busca exploratória, esses resultados individuais dos blocos devem ser avaliados e relacionados com os resultados anteriores. Esse relacionamento pode evitar retrabalho ao avançar etapas no Framework. A Tabela 2 apresenta uma explicação mais geral, com as entradas dos processos, as atividades dos processos e a saída de cada bloco do Framework.

4.1.1 Coleta de Dados - A

A coleta e processamento inicial dos dados (bloco azul) deve ser considerada uma das etapas mais importantes do problema, visto que aqui se concentra a definição do problema. Nesta etapa devem ser coletadas as apresentações definidos quais tipos de dados vão ser coletados. Ainda, com os dados em mãos, devem ser classificados em posturas/gestos/ações/voz, de modo a identificar o que o estudante fez durante as apresentações. A qualidade desses dados, e por conseguinte a capacidade de realmente representar as apresentações dos alunos, vai definir o resultado do problema e a descoberta dos padrões.

Nesta etapa os dados são organizados de maneira que possam ser feitas várias análises posteriores. Assim eles devem estar dispostos de maneira que preservem certos aspectos das apresentações. Exemplos disso são as sequências das posturas, que podem ser representadas com o tempo (instante ou intervalo) na apresentação que ocorreu, ainda deve-se preservar quais alunos realizaram as apresentações, para uma posterior identificação. Isso pode ser feito com índices que representariam os

Tabela 2 – Detalhes dos processos do Framework agrupados. Fonte: Elaborado pelo autor.

Block	Process
A	<p>Input: Presentations of students.</p> <p>Activity:</p> <ol style="list-style-type: none"> (1) Capture data from oral presentations; (2) Classification of data in body/voice/action; (3) Tabulate the data; (4) Generate the database containing the observations with their times added to each attribute; <p>Output: Database containing the students' time intervals with respective body posture/voice.</p>
B	<p>Input: Database containing the students' time intervals.</p> <p>Activity:</p> <ol style="list-style-type: none"> (1) Data visualization comparing the databases; (2) Average of the values of each attribute; (3) Inferential test comparing the two or more databases. <p>Output: Information graphics; Table comparing the most present attributes; Inferential comparison between databases and evaluated presentations.</p>
C	<p>Input: Database with observations.</p> <p>Activity:</p> <ol style="list-style-type: none"> (1) Perform clustering techniques; (2) Save all results to evaluation; (3) For each cluster value, run a cluster evaluation algorithm; (4) Visually evaluate the results if possible; (5) Create/use analytics methods to choose the best value of clusters to the problem; (6) Critical inspection on data results to find patterns. <p>Output: Decision of the best cluster value for the problem; Patterns that clustering method allows to find; plots showing the patterns and characterization of clusters.</p>
D	<p>Input: Initial database.</p> <p>Activity:</p> <ol style="list-style-type: none"> (1) Transform the initial database into a sequential database; (2) Transform the data to enter the tool; (3) Test minimum Support values; (4) Create visual methods to discover the patterns; (5) Correlation the sequences results with the other data results. <p>Output: More frequent sequences given a minimum support value; Relation between the other results.</p>

alunos.

A base de dados para o restante das análises dos resultados também deve ser gerada. As observações (estudantes com seus respectivos tempos ou outro tipo de dado que represente o problema em cada um dos atributos coletados). É importante ressaltar que uma quantidade maior de dados pode permitir realizar conexões com resultados dos blocos e métodos, que por suas próprias características são diferentes. Dados como notas e slides da apresentação podem ser usados em análises posteriores.

4.1.2 Análise Estatística - B

A etapa de análise estatística, identificada pelo bloco laranja, é a primeira parte dos resultados propriamente ditos do Framework. Nesta etapa, uma análise com estatística descritiva é feita. O objetivo principal dessa análise descritiva é compreender os dados de forma geral. Medidas de dispersão de dados podem ser usadas aliadas a gráficos que são gerados de modo a evidenciar algum comportamento relevante dos dados. As linhas tracejadas brancas, de certa forma, representam o fluxo, ainda que ele não seja um processo contínuo e muitas vezes as etapas devem ser revistas ao fim de identificar falhas de algum procedimento.

Nesta etapa também é gerada a análise com estatística inferencial. Este método testa duas ou mais populações diferentes que neste caso seriam as apresentações. O teste pode indicar semelhança estatística entre as apresentações e anos de avaliação. Ou ainda apontar diferença estatística. Nos dois casos, resultados importantes para entender o processo de aprendizagem dos estudantes. Esses resultados podem ser usados para definir estratégias de ensino com base nos resultados de diferentes anos de apresentações.

A comparação pode se dar entre anos diferentes de mesmo curso, ou ainda pode ser entre cursos diferentes de mesmo ano, não existindo limitações nesse sentido. Ressalta-se para atentar na realização de testes de normalidade anteriores a realização do teste inferencial, de modo a confirmar a distribuição das apresentações.

4.1.3 Clustering - C

Seguindo a linha tracejada chega-se no bloco C, que representa o aspecto dos algoritmos de clustering no framework. Um dos objetivos dessa abordagem é poder segmentar os estudantes em diferentes grupos com características bem definidas para uma avaliação granular dos comportamentos. Podendo assim diferenciar as apresentações e encontrar e relacionar resultados.

É difícil identificar o número de grupos em um algoritmo de clustering, logo, neste framework sugere-se o teste de diferentes valores de clusters para o problema.

Métodos de avaliação de clusters então devem ser usados para identificar o melhor número de grupos dentro das apresentações.

Uma busca exploratória nos resultados do algoritmo de clusterização deve ser realizada, respeitando o tipo de resultado do algoritmo escolhido. Essa busca deve focar na descoberta de características e comportamentos nos resultados que caracterizem os grupos formados pelo algoritmo. Em geral, algoritmos de clusterização usam alguma regra para formar os grupos. Essa descoberta pode ser relacionada com resultados outros resultados, como as inferências feitas no passo anterior.

Os blocos menores representam processos menores mas que devem ser realizados ao utilizar o Framework. A busca exploratória faz parte de todo processo do Framework, contudo aqui ela tem destaque visto que a formação dos clusters contém informação do comportamento dos estudantes. A visualização de dados podem esclarecer padrões que são difíceis de observar ao olhar uma tabela, por exemplo.

4.1.4 Sequential Pattern Mining - D

O *Sequential Pattern Mining* surge como uma alternativa e complementação dos resultados encontrados nas análises anteriores. A medida que existe uma base sequencial com os intervalos de cada atributo que o estudante passou, essa base pode ser usada para encontrar padrões sequenciais nos dados.

Assim, o resultado do algoritmo são as sequências mais presentes dado um valor mínimo de suporte. Essas sequências podem ser uma maneira interessante de avaliar o comportamento do estudante ao longo da apresentação, ela pode indicar um estudante com um comportamento de menor participação na apresentação, por exemplo.

Além das sequências, essa etapa deve conter visualizações das sequências que não em forma escrita. Isso deve ser feito para facilitar a compressão dos resultados encontrados, visto que eles podem ser difíceis de relacionar e identificar aspectos interessantes. Ainda, a importância de existirem outros tipos de dados, como as notas, uma vez que estas podem ser relacionadas com as sequências encontradas.

4.2 ESTUDO DE CASO

4.2.1 Ambiente de Aprendizagem das Apresentações

Esta seção apresenta o estudo de caso realizado. A descrição do caso, do dispositivo de coleta de dados, da coleta de dados, da ferramenta de organização dos dados e uma descrição dos dados são apresentadas aqui.

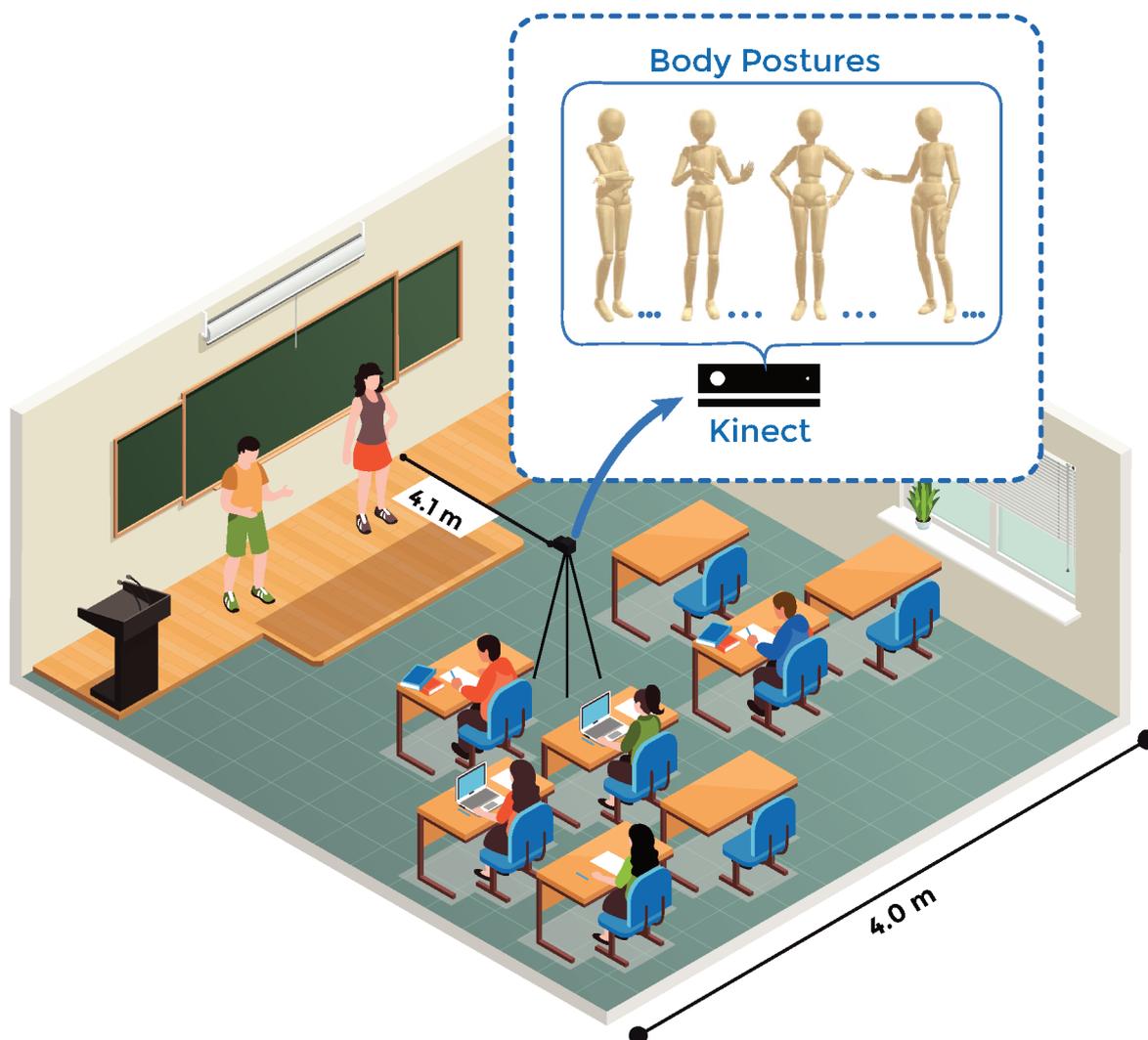


Figura 5 – Configuração do ambiente. Fonte:(ROQUE, F. *et al.*, 2019b)

4.2.1.1 Descrição do Caso

O contexto de aprendizagem do presente estudo foi o curso de “Introdução à engenharia” oferecido a estudantes de Engenharia da Computação da Universidade de Valparaíso, no Chile, durante os anos 2017 e 2018. Durante o curso, os alunos realizaram três apresentações em grupos orais sobre três tópicos previamente predefinidos no curso básico, como será mostrado na seção 4.2.1.3. Um esquema da configuração do ambiente de aprendizado é apresentado na Figura 5.

4.2.1.2 Microsoft Kinect

A captura de dados para este trabalho foi realizada pelo Microsoft Kinect. A ferramenta é capaz de capturar até 6 pessoas simultaneamente. O Microsoft Kinect provou ser eficiente na extração dos padrões de postura e fala necessários para os experimentos. Suas vantagens estão relacionadas ao número de sensores presentes

no dispositivo. Ele contém um sensor de profundidade, uma câmera colorida e um conjunto completo de microfones que fornece recursos de captura de movimento em 3D de corpo inteiro, reconhecimento facial e reconhecimento de voz (ZHANG, 2012).

Foi utilizado o rastreamento esquelético para classificar as posições dos alunos e a detecção de voz. Com isso, foi possível representar o corpo humano como um número de articulações. Essas articulações são partes do corpo: braços, pescoço, mãos e cabeça, por exemplo. A Figura 6 mostra duas pessoas e suas representações pelo rastreamento esquelético. Essas representações do rastreamento esquelético foram usadas para classificar ações pré-definidas e treinadas e, assim, montar os bancos de dados utilizados. O tempo gasto em cada postura / ação também foi armazenado.

4.2.1.3 Procedimentos e Coleta de Dados

Os dados capturados correspondem a apresentações orais no curso de Introdução à Engenharia. O curso é oferecido no primeiro semestre para os alunos do curso de Engenharia da Computação (na Universidad de Valparaíso, Chile). No curso, os alunos são avaliados em três apresentações orais avaliadas, associadas a 1) desenvolvimento web, 2) microcontroladores e 3) banco de dados (com um mês entre cada apresentação). Cada apresentação foi realizada por grupos de 2 alunos, com no máximo 5 minutos por grupo. Os slides do PowerPoint podem suportar todas as apresentações dos alunos. Todas as apresentações foram capturadas usando o software Microsoft Kinect e a aplicação Leikelen (MUNOZ *et al.*, 2018). Os alunos tiveram que assinar um consentimento informado no início do curso. Além disso, a metodologia foi aprovada pelo Comitê de Ética em Pesquisa com Seres Humanos da Faculdade de Medicina da Universidade de Valparaíso.

4.2.1.4 Ferramenta Leikelen

Foi usado o software Leikelen (MUNOZ *et al.*, 2018) para coleta e processamento de dados (consulte a Figura 6). O software permite detectar, armazenar e visualizar posturas corporais das pessoas gravadas. O Leikelen tem opções para adicionar posturas personalizadas, exportar e importar cenas e adicionar um visualizador das posturas detectadas, juntamente com uma linha do tempo. Por fim, o Leikelen também permite exportar os dados a serem visualizados usando outras ferramentas de mineração de dados.

A aplicação Leikelen usa o Microsoft Kinect para capturar as posições do corpo e o algoritmo ADABOOST (FREUND; SCHAPIRE, 1995) para classificar as posturas. O software possui 10 posturas corporais pré-definidas e validadas, juntamente com outros tipos de métricas, como 1) distância proxêmica (distância entre duas ou mais pessoas) (HALL *et al.*, s.d.), 2) se a pessoa está olhando para o público, 3) a inclinação do corpo e 4) se a pessoa está falando. Para este estudo, as posturas utilizadas foram

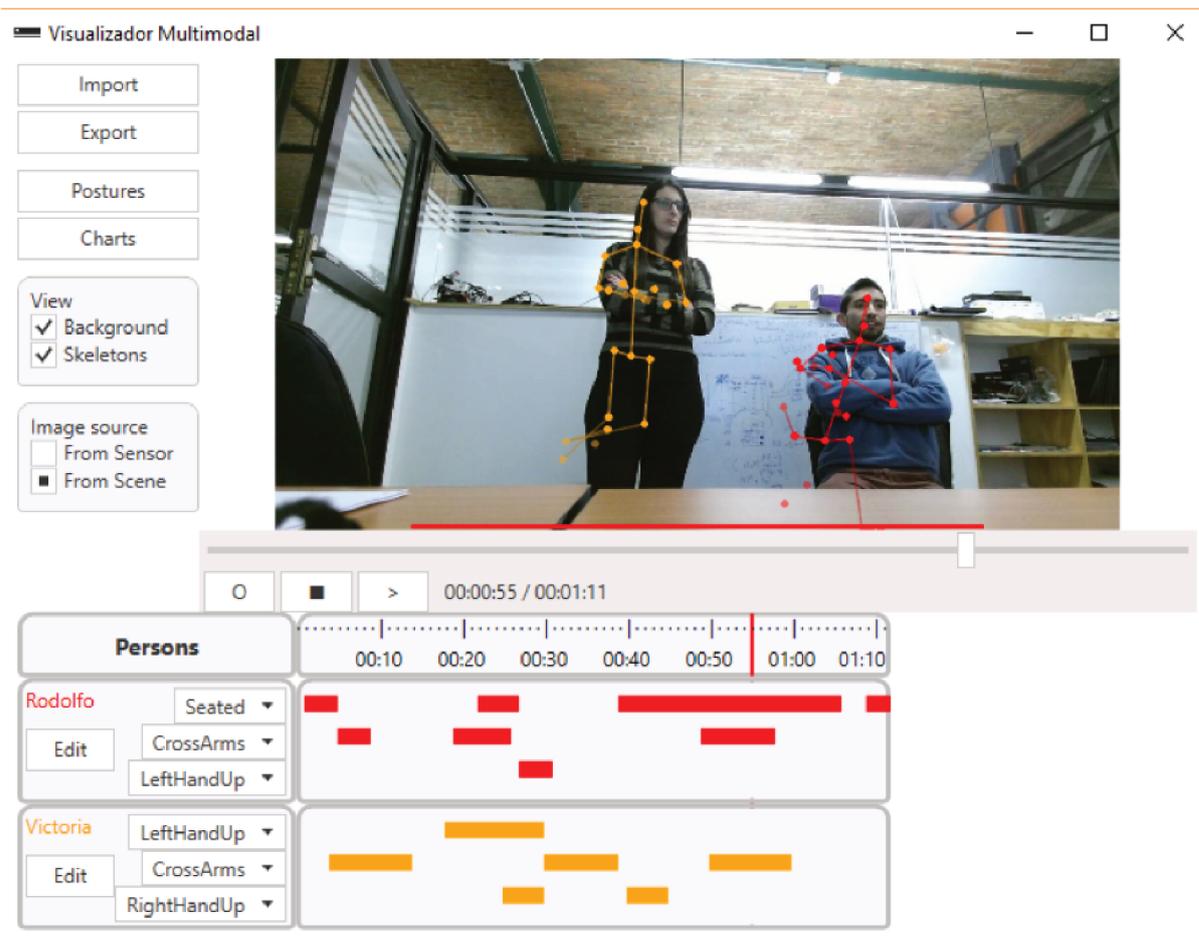


Figura 6 – Tela do aplicativo Leikelen mostrando o modelo de *skeleton* de duas pessoas.

incorporadas a partir de (MUNOZ *et al.*, 2018), baseando-se principalmente em (CHEN *et al.*, 2015; ECHEVERRIA *et al.*, 2014).

4.2.1.5 Características Coletadas (Atributos)

Os atributos utilizados no presente trabalho são mostrados na Tabela 3. Os dados brutos são compostos por uma sequência de períodos e a respectiva ação / postura do aluno capturada pelo sistema durante esse período. Por exemplo, um aluno pode passar um certo período com os braços cruzados e, em seguida, deixar os braços para baixo por um período muito curto de tempo e começar a apontar para os slides, finalmente retornando à posição com os braços cruzados. Para esta sequência, obtemos quatro registros compostos pelas ações / posturas capturadas (Braços Cruzados, Mãos Baixas, Apontando e Braços Cruzados), juntamente com o período em que cada ação / postura durou (MUNOZ *et al.*, 2018).

Foram coletados um total de 6 conjuntos de dados de alunos diferentes durante os anos de 2017 e 2018 (três apresentações por ano), como pode ser visto na Tabela

Tabela 3 – Descrição de cada atributo das bases de dados.

Atributos	Descrição
Braços Cruzados	O apresentador cruzou os dois braços.
Cabeça Baixa	A inclinação do apresentador é maior que 0,333, com -1 inclinado para trás e 1 inclinado para frente.
Ereto	A inclinação do apresentador está entre -0,333 e 0,333, com -1 inclinado para trás e 1 inclinado para frente.
Olhando o Público	O apresentador está olhando para o público.
Mão na Face	O apresentador tem uma mão no queixo.
Mãos Abertas	O apresentador está explicando com as duas mãos (ambas as mãos com os braços cruzados).
Mãos Baixas	O apresentador está de mãos dadas
Uma Mão	O apresentador está explicando com uma mão para baixo e a outra dobrou em uma posição explicativa.
Mão na Cintura	O apresentador está com as mãos na cintura.
Mão na Cabeça	O apresentador tem uma mão na nuca.
Apontando	O apresentador está apontando com uma mão.
Fala	A voz do apresentador é detectada.

Tabela 4 – Número de observações de cada base de dados.

Ano	Apresentação	Número de Estudantes
	P1	40
2017	P2	22
	P3	23
	P1	59
2018	P2	45
	P3	34

4.

4.2.1.6 Descrição dos Dados

Normalmente, uma observação E é representada por uma coleção de valores de atributos (características) $(x_1; x_2; \dots; x_n)$, onde x_i é um valor de um atributo X_i . Para esse problema, um exemplo é um aluno em uma apresentação. As observações deste problema foram obtidas através da soma dos intervalos de tempo de cada atributo, divididos pelo tempo total de apresentação, portanto, o máximo possível para uma determinada observação de um atributo é 1.

A Tabela 5 mostra os valores médios de cada atributo em cada conjunto de dados. Os valores são normalizados (variando de 0 a 1) e foram obtidos da soma dos valores do atributo fornecido na base de dados dividido pelo número de observações

Tabela 5 – Valor médio para cada atributo.

	2017			2018		
	P1	P2	P3	P1	P2	P3
Braços Cruzados	0,249	0,247	0,281	0,207	0,188	0,243
Cabeça Baixa	0	0,005	0	0,012	0	0
Ereto	0,930	0,935	0,939	0,808	0,901	0,856
Olhando o Público	0,493	0,529	0,496	0,473	0,468	0,365
Mãos na Face	0,074	0,061	0,053	0,112	0,072	0,084
Mãos Abertas	0,201	0,209	0,249	0,148	0,170	0,157
Mãos Baixas	0,506	0,447	0,427	0,465	0,518	0,442
Uma Mão	0,132	0,102	0,100	0,111	0,148	0,145
Mãos na Cintura	0,034	0,048	0,026	0,050	0,067	0,041
Apontando	0,105	0,099	0,099	0,230	0,183	0,180
Mão na cabeça	0,001	0,003	0,002	0,009	0,003	0,004
Fala	0,212	0,249	0,222	0,107	0,169	0,124

no conjunto de dados, como pode ser visto na Equação 4,

$$\bar{A} = \frac{\sum_{j=1}^N A_j}{N} \quad (4)$$

em que \bar{A} é o valor médio de um atributo, N é o número de observações de uma apresentação e A_j é o valor de um atributo na base de dados. Como pode ser visto na Tabela 5, alguns atributos são mais prevalentes que outros. O atributo que mais ocorre nos conjuntos de dados é a postura *Ereto*. Nas três apresentações de 2017, esse atributo ocorreu por mais de 90% do tempo total, enquanto no ano 2018, o atributo ocorreu por cerca de 80% do tempo total. As posturas *Olhando o público* e *Mãos Baixas* também ocorrem com bastante frequência e se destacam das demais. As ocorrências mais altas para esses dois primeiros atributos (*Ereto* e *Olhando o público*) são esperadas, pois durante a apresentação é bastante normal que os alunos estejam em pé e constantemente olhando para o público. Os principais atributos de 5 predominantes em cada apresentação são apresentados na Tabela 6. Para 2017, os cinco principais atributos que ocorreram mais são quase os mesmos, com uma pequena diferença na classificação (exceto pelo atributo *Mãos Abertas* em *P3*). O mesmo vale para 2018, onde a diferença entre as apresentações é para o atributo *Braços Cruzados* em *P3*.

Conforme mostrado na Tabela 5, alguns atributos não variam significativamente entre as apresentações. Por exemplo, *Braços Cruzados* e *Mãos Baixas* apresentam valores muito semelhantes nas apresentações dos dois anos. Por outro lado, alguns atributos variam entre apresentações e em comparação entre os anos. Essas diferenças são melhor analisadas na seção 5, na qual compara-se as características dos conjuntos de dados e discute-se os padrões mais importantes encontrados entre eles.

Tabela 6 – 5 atributos mais presentes de cada base de dados.

Ano	P1	P2	P3
2017	Ereto	Ereto	Ereto
	Mãos Baixas	Olhando público	Olhando público
	Olhando público	Mãos Baixas	Mãos Baixas
	Braços Cruzados	Fala	Braços Cruzados
2018	Ereto	Ereto	Ereto
	Mãos Baixas	Olhando público	Mãos Baixas
	Olhando público	Mãos Baixas	Olhando público
	Braços Cruzados	Braços Cruzados	Apontando

5 RESULTADOS - ESTUDO DE CASO

Nesta seção são apresentados os resultados do estudo de caso realizado com os dados. Na seção é apresentado a comparação dos atributos entre os cursos, a etapa de geração dos grupos (clusters), a comparação entre os grupos, a análise de padrões sequenciais, a discussão e implicações dos resultados.

5.1 COMPARAÇÃO DOS ATRIBUTOS ENTRE OS CURSOS

Uma maneira de comparar os anos avaliados no estudo é gerar gráficos de tendência entre os anos de 2017 e 2018. Ao analisar os gráficos, é possível observar esses atributos com uma tendência semelhante (onde os valores dos atributos de um ano seguem o do outro ano), em outros atributos esse comportamento não é observado e os gráficos têm comportamentos aleatórios. Ainda outro comportamento, em que, em algum momento (ao longo das apresentações), os valores dos atributos são muito próximos, podem ser vistos no atributo *MãosNaFace*. Estes resultados são mostrados na Figura 7.

O primeiro grupo de atributos (valores similares de evolução ao longo das apresentações) tem uma semelhança muito clara e pode ser observado no gráfico variável de fala por exemplo, as Figuras 7(a) e 7(b). Nesse grupo, o comportamento é o mesmo nas três apresentações do semestre; se um valor da variável cai em um ano na segunda apresentação ela também cai na mesma apresentação do outro ano. Nota-se que essa diminuição ou aumento no valor do atributo varia entre anos.

O outro grupo de atributos são os que possuem diferenças de comportamento ao longo das apresentações, como a variável *Cabeça Baixa*, Figura 7(c). Os valores das porcentagens de tempo são diferentes entre os anos e, aparentemente, não há similaridade de comportamento entre os anos avaliados. O que se percebe é que os valores são próximos entre os anos, mesmo nesse grupo de atributos em que a tendência mostrada anteriormente não é observada. Ainda é possível observar atributos em que, em algum momento das apresentações, obtiveram uma tendência semelhante, mas que divergiu em outro ponto do semestre, como no atributo *Mãos Baixas*, Figura 7(d).

Para melhor comparar as bases de dados, foram utilizadas algumas métricas estatísticas inferenciais. Primeiro, foi analisada a hipótese nula de que os atributos são provenientes de uma distribuição normal padrão (LILLIEFORS, 1967). Os resultados mostraram que alguns atributos das apresentações de 2018 têm distribuição normal. Assim, foi necessário o uso de um método não paramétrico para testar a hipótese nula de que dados das duas bases de dados diferentes são amostras de distribuições contínuas com medianas iguais.

O teste *Wilcoxon Rank Sum* (GIBBONS; CHAKRABORTI, 2011) é um teste não

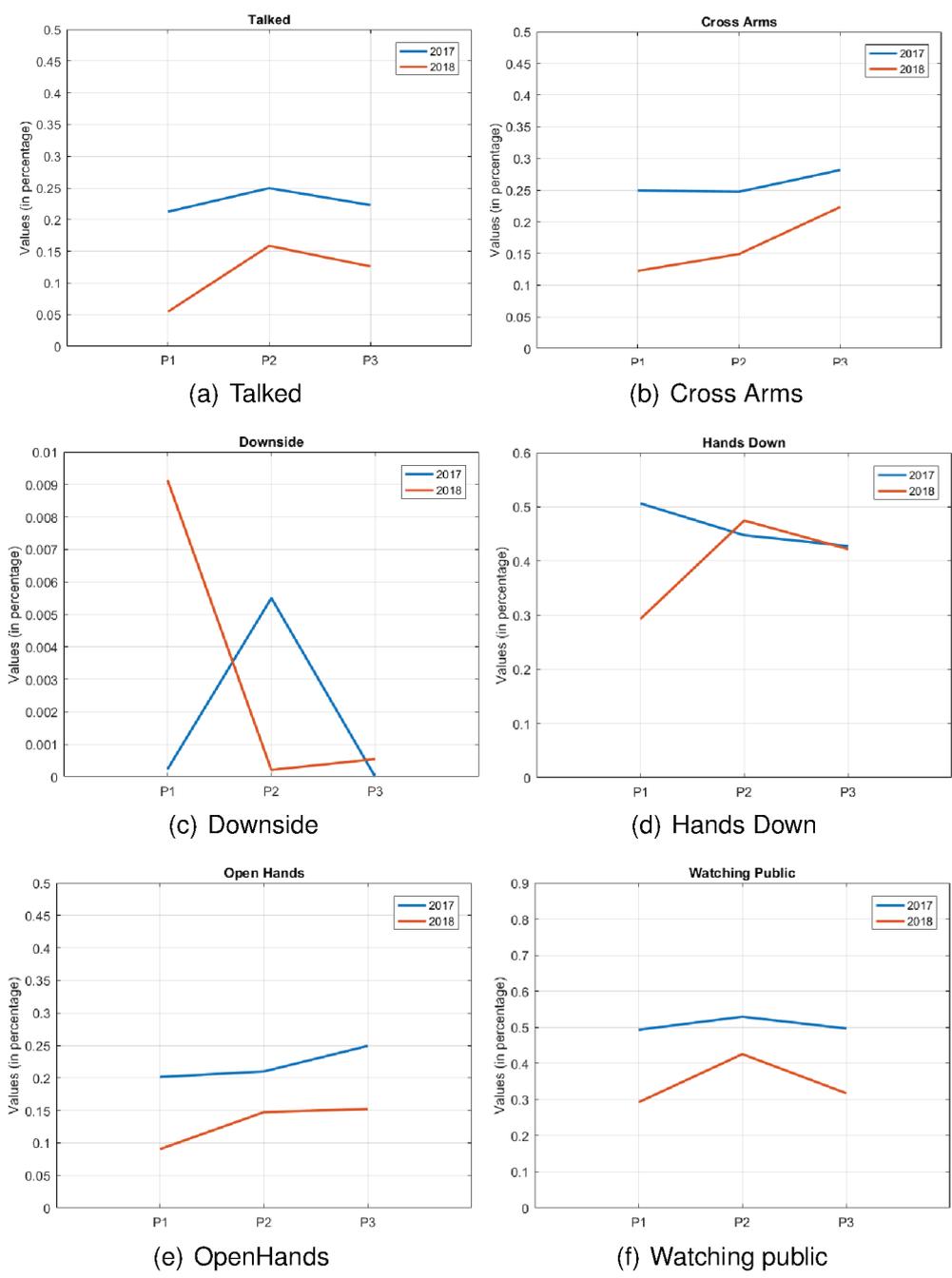


Figura 7 – Gráficos de tendência comparando os anos. Fonte: Adaptado de (ROQUE, F. *et al.*, 2019b).

paramétrico para duas populações, quando as amostras são independentes. O teste *Wilcoxon Rank Sum* é equivalente ao teste U de Mann-Whitney. Para este trabalho, o valor p foi utilizado para indicar a igualdade das duas medianas da população, com nível de significância de 5%. Valores altos de p indicam que não há evidência suficiente para rejeitar a hipótese nula. Valores baixos de p indicam o contrário.

A Tabela 7 apresenta os testes para cada par de duas apresentações (bases de dados), combinando todas as possibilidades, totalizando 15 comparações entre as bases de dados. Para facilitar a análise, foram excluídas as comparações entre as apresentações de diferentes momentos, de diferentes anos. Consequentemente, a tabela mostra apenas comparações entre os mesmos anos e comparações entre os mesmos momentos de anos diferentes. Para facilitar a leitura da tabela, os valores considerados pequenos foram arredondados para 0. Como pode ser visto, algumas combinações obtiveram baixos valores de p . Isso indica que nesses atributos, as populações têm medianas estatisticamente diferentes. Os valores em negrito representam os resultados com medianas estatisticamente diferentes.

No ano de 2017, as medianas obtiveram valores estatisticamente semelhantes. Em comparação com a primeira apresentação com a segunda, por exemplo, apenas 1 atributo foi estatisticamente diferente. Para a comparação entre a primeira e a terceira apresentações, 0 atributo estatisticamente diferente. No entanto, na última comparação, entre a segunda e a terceira apresentações, 2 atributos testados como diferentes. Ainda, para o ano de 2017, apenas *Cabeça Baixa* dois testes não apresentaram semelhanças entre as medianas.

Já para as apresentações de 2018, mais atributos apresentaram diferenças significativas. A primeira comparação, entre as apresentações um e dois, obteve 3 atributos considerados estatisticamente diferentes; em relação a 2017, esse resultado pode ser considerado semelhante, pois possuímos 12 atributos para avaliar. A comparação entre a primeira e a terceira apresentação teve 1 atributo com diferentes medianas. Na última comparação de 2018, entre a segunda e a terceira apresentações, registrou-se 0 medianas estatisticamente diferentes. Em 2018, *Cabeça Baixa* dois testes novamente apresentaram uma mediana estatisticamente diferente.

Comparando os dois anos, com as apresentações nos mesmos momentos (momentos em sequência), é possível observar que durante as apresentações as diferenças de atributos tenham sido reduzidas de 6 para 1. Na comparação com a primeira apresentação dos dois anos (2017 e 2018), houve mais diferenças estatísticas dos atributos, com 6. As primeiras apresentações são aquelas que contêm mais observações nas bases de dados, o que pode gerar alguma variação.

Como mencionado, na comparação das primeiras apresentações, tem-se um total de 6 atributos estatisticamente diferentes. Podemos visualizar que, na comparação entre as segundas apresentações de cada ano, o valor cai de 6 para 3, indicando que

Tabela 7 – Testes inferenciais entre as bases de dados. Fonte: (ROQUE, F. *et al.*, 2019b)

	2017			2018			2017X2018		
	P1xP2	P1xP3	P2xP3	P1xP2	P1xP3	P2xP3	P1xP1	P2xP2	P3xP3
Cross Arms	0,900	0,622	0,741	0,338	0,590	0,310	0,309	0,197	0,459
Downside	0,046	0,290	0,017	0	0,001	0,744	0	0,023	0,251
Straight	0,213	0,926	0,369	0,006	0,084	0,425	0,003	0,607	0,510
Talked	0,361	0,954	0,495	0,060	0,493	0,282	0	0,030	0,065
Watching public	0,888	0,684	0,829	0,624	0,065	0,590	0,917	0,376	0,168
Hand On Face	0,842	0,357	0,357	0,669	0,804	0,831	0,509	0,952	0,317
Open Hands	0,653	0,082	0,346	0,321	0,734	0,666	0,025	0,179	0,007
Hands Down	0,327	0,216	0,657	0,259	0,876	0,164	0,283	0,248	0,763
One Hand	0,168	0,497	0,593	0,029	0,102	0,909	0,493	0,017	0,102
Hand On Hip	0,906	0,054	0,108	0,952	0,105	0,150	0,723	0,784	0,604
Point	0,900	0,858	0,691	0,309	0,352	0,964	0,002	0,062	0,086
Hand On Head	0,264	0,138	0,029	0,291	0,050	0,384	0,030	0,826	0,189
Total	1	0	2	3	1	0	6	3	1

os alunos apresentam (em relação aos atributos avaliados) maior semelhança entre os anos. Finalmente, em comparação com as últimas apresentações, vemos que os atributos estatisticamente diferentes caem para 1. Uma nova queda no número total de atributos diferentes entre os anos, o que pode indicar que, no final, os alunos apresentaram um padrão mais semelhante com respeito aos dois anos.

5.2 GRUPOS (CLUSTERIZAÇÃO)

Para avaliar o contexto, onde os padrões de apresentação são encontrados, alguns algoritmos foram utilizados para explorar os dados. O algoritmo *K-means* agrupa exemplos semelhantes ao mesmo conjunto, de acordo com suas características. Dessa forma, os valores dos centroides podem ser analisados. Mesmo assim, em muitos casos, é difícil e, várias vezes, não se sabe quantos grupos o problema possui. O algoritmo que gera as *Silhouettes* ajuda nesse sentido, auxiliando visualmente na identificação de casos em que o *K-means* está errando em alguns exemplos.

5.2.1 Os Grupos (Clusters)

Foram avaliados 6 valores diferentes de k para cada uma das bases de dados. Para cada valor de k , os centroides e os índices das observações (o grupo que a observações pertence) foram salvos. A avaliação de *Silhouettes* ajuda a encontrar o melhor valor de k para cada base de dados, uma vez que o número de grupos não é conhecido, nem que tipos de comportamento eles apresentam. O resultado de todas as avaliações para cada valor de k é mostrado na Figura 8 e Figura 9.

O número de observações em cada apresentação é diferente, conforme explicado anteriormente na Tabela 4. Dessa maneira, o melhor número de k pode variar entre as bases de dados avaliadas. Na Figura 9, a primeira apresentação do segundo ano tem em todos os valores avaliados de k um grande grupo. Isso pode ser considerado um comportamento diferenciado em relação às outras duas apresentações, que conforme o aumento de k as observações estão se separando entre os grupos (clusters).

Observando a Figura 8 e a Figura 9, os valores iniciais de k têm poucas observações com valores negativos de *Silhouettes* (indicando que a observação não pertence ao cluster). Isso sugere que o melhor número de grupos está nos primeiros valores de k . Para o primeiro ano de avaliação (2017), pode-se notar também que os melhores agrupamentos estão nos dois primeiros valores de k . Após $k = 3$, o algoritmo começa a atribuir grupos com apenas uma observação. Existem também muitos valores negativos de *Silhouettes*.

Como mencionado, nem sempre é fácil encontrar o número certo de grupos para um problema de aprendizado não supervisionado. Neste trabalho foi utilizada a técnica visual de *Silhouettes* e os valores médios das *Silhouettes* também foram utilizados na avaliação. Os centroides de cada grupo também foram avaliados, para este trabalho buscou-se uma clara separação entre os centroides dos atributos de cada grupo, em um determinada base de dados. A Tabela 8 mostra os valores médios das *Silhouettes*. Nesta etapa do trabalho, foram encontrados os melhores valores de k para cada ano, para posteriormente serem analisados em cada uma das bases de dados (apresentações) agrupadas.

Nas apresentações do ano de 2017, com base na tabela de médias e na análise visual, decidiu-se analisar os valores de $k = 3$. É importante deixar claro que esse valor de k também foi escolhido após uma análise inicial dos centroides. Isso foi feito com o intuito de evitar os altos valores que as *Silhouettes* também apresentavam, com pouca separação entre os valores dos centroides dos grupos. Portanto, esse alto valor médio pode parecer uma boa opção de valor para os clusters. Um exemplo desse comportamento pode ser visto no valor de $k = 4$ para a apresentação 3, no ano de 2017, Figura 8(i).

Para o ano de 2018, houve uma variação maior nos valores de k . O valor de

Tabela 8 – Média dos valores das *Silhouettes*

	2017			2018		
	P1	P2	P3	P1	P2	P3
k=2	0,417	0,478	0,437	0,280	0,387	0,380
k=3	0,310	0,382	0,443	0,259	0,283	0,361
k=4	0,227	0,192	0,310	0,237	0,247	0,190
k=5	0,173	0,286	0,247	0,274	0,255	0,204
k=6	0,193	0,232	0,208	0,262	0,230	0,268
k=7	0,217	0,303	0,253	0,165	0,163	0,215

$k = 3$ foi escolhido novamente. Este valor ($k = 3$) apresenta um bom agrupamento, como pode ser visto nas Figuras 9(d), 9(e) e 9(f). O valor médio das *Silhouettes* não apresenta valores tão altos, mas para comparar as apresentações com o ano de 2017, é importante escolher o mesmo valor de k para os dois anos.

Para este trabalho, o valor de $k = 3$ foi escolhido, pois apresenta bons valores médios de *Silhouettes* e também apresenta uma formação homogênea de clusters, como pode ser visto nas *Silhouettes*. Outro motivo que levou a essa escolha foi o fato de existirem três padrões distintos, quando se olha para os centroides. Escolher, por exemplo, $k = 2$, como o número de clusters, faria com que um grupo de alunos estivesse dentro de outro cluster. A análise de centroides para $k = 3$ mostra que existem diferenças sutis, mas importantes, no contexto de aprendizado que precisam ser exploradas.

5.3 COMPARAÇÃO ENTRE OS CLUSTERS - ANÁLISE DOS CENTROIDES

Esta seção apresentará uma análise dos valores dos centroides, entre as diferentes apresentações nos momentos determinados. Nesta etapa, como mencionado anteriormente, os centroides já gerados foram avaliados. O objetivo desta etapa de avaliação é encontrar algo relacionado à evolução ou comportamento de cada cluster gerado com base em seus atributos. A Figura 10 mostra alguns atributos (8 atributos) em um gráfico polar. Optou-se por não plotar os outros quatro atributos, considerando que eles não estão presentes no gráfico, pois seus valores de centroide são muito pequenos.

Alguns atributos se destacam mais e outros menos, em certos clusters formados. O trabalho de (CAVANAGH *et al.*, 2014) apresenta algumas características de excelentes desempenhos durante uma apresentação, e outras características de desempenhos chamados pobres. Essas características podem estar relacionadas aos atributos avaliados no presente trabalho, deixando claro o comportamento dos clusters.

As características consideradas ruins estão relacionadas à linguagem corporal retraída, como: posicionamento defensivo do braço (braços cruzados, mãos nos bol-

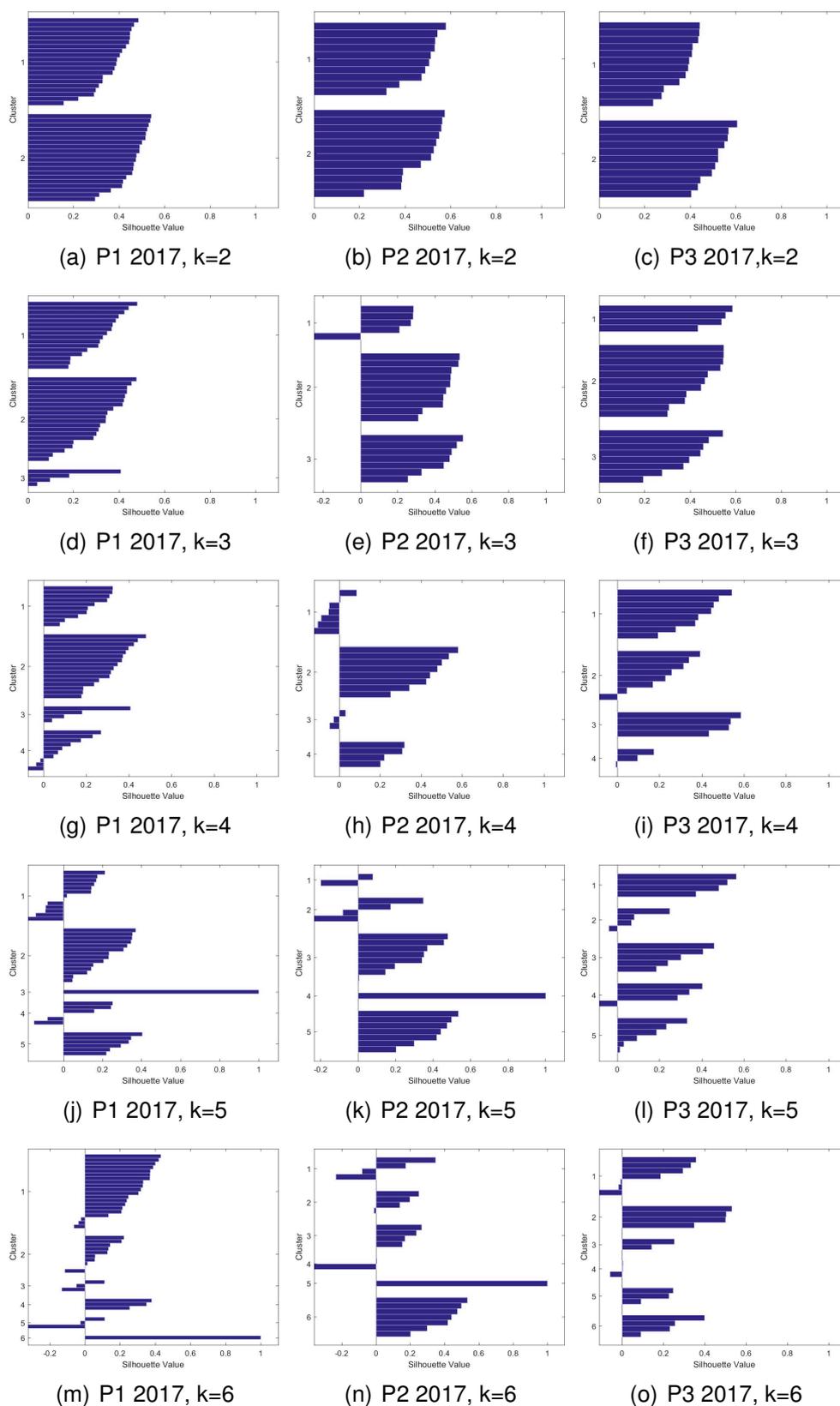


Figura 8 – Gráficos das *Silhouettes* para três diferentes momentos em 2017. Fonte: Adaptado de (ROQUE, F. *et al.*, 2019b)

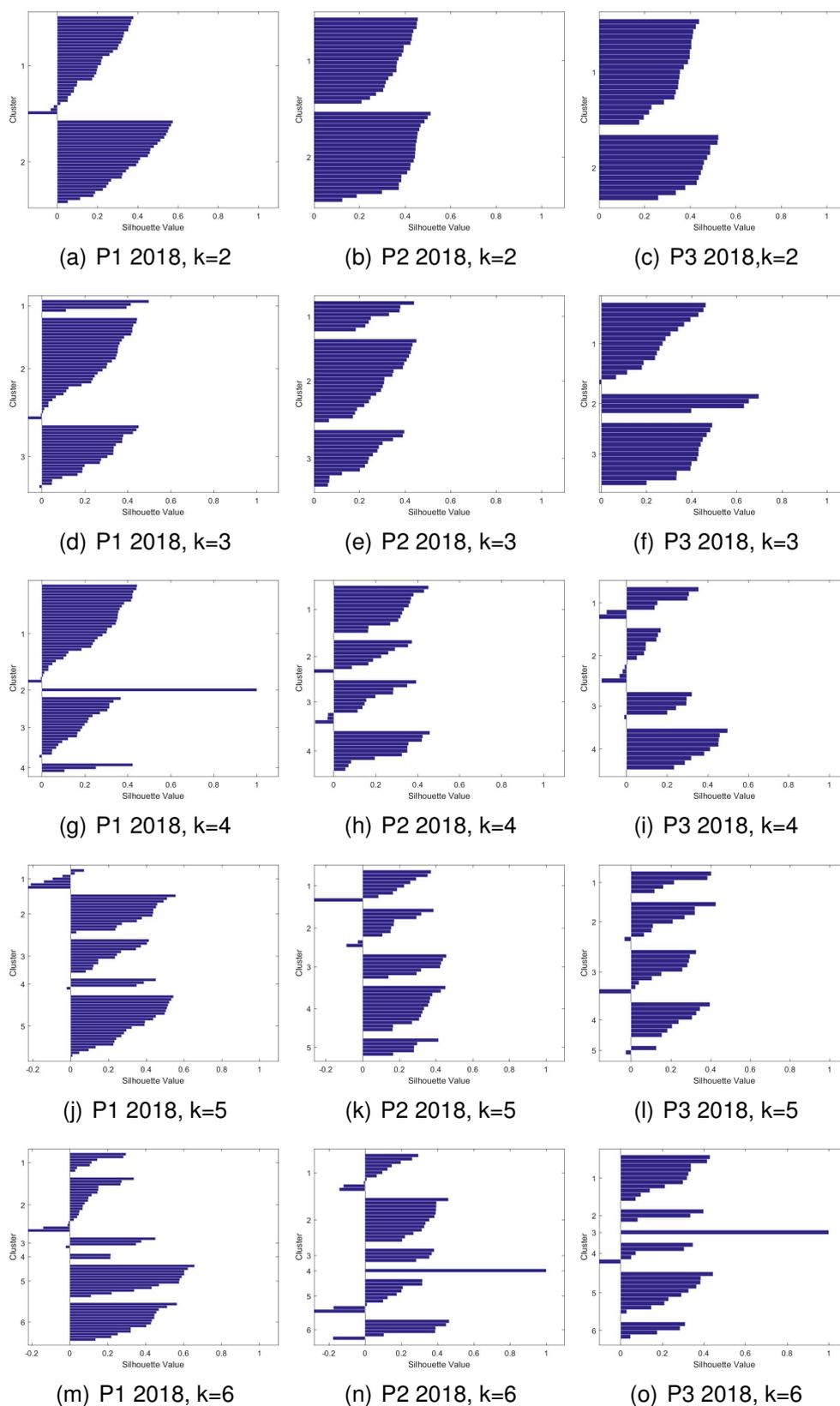


Figura 9 – Gráficos das *Silhouettes* para três diferentes momentos em 2017. Fonte: Adaptado de (ROQUE, F. et al., 2019b)

sos), postura retraída e cabeça para baixo. As características consideradas excelentes apresentações pelos autores são: postura corporal aberta, gestos com as mãos para enfatizar pontos ou transmitir significado e contato visual inclusivo. No trabalho são apontadas outras características, as que são mencionadas aqui são as que mais se relacionam com os atributos avaliados no presente trabalho. Também pode-se associar a fala a algo positivo em uma apresentação, pois não falar indica, em muitas situações, que o aluno explicou menos sobre o assunto.

5.3.1 Apresentações de 2017

Na primeira apresentação de 2017, alguns atributos diferem em relação aos clusters. Nesta análise, alguns deles são avaliados, sendo escolhidos devido à sua importância no comportamento dos clusters. O atributo *Braços Cruzados*, por exemplo, apresenta uma diferença significativa entre os três clusters. Este tipo de atributo é altamente relevante quando os alunos se comunicam. Isso ocorre porque, por exemplo, braços cruzados são um típico gesto defensivo (LUCAS, 2011). Existe um grupo, grupo A, de observações que possui mais presente essa postura, outro que é menor, grupo C, e ainda outro grupo de observações (grupo B) que fica no meio entre os outros dois. Para esta apresentação, ainda tem-se o atributo *Fala*. Apresenta a mesma característica do atributo anterior, onde existem três “ marcações ” distintas nos três clusters. Uma com um valor mais alto, outra como um meio termo e uma com valor menor.

O atributo *Olhando o Público* se comporta de maneira diferente dos outros dois. Um cluster tem um entendimento diferente em comparação com os outros dois, que têm quase 100% do tempo da apresentação. O mesmo padrão, não para os mesmos clusters, é encontrado no atributo *Apontando*. Dois clusters (A e C) têm valores baixos e um (grupo B) possui valores mais altos. A maioria dos outros atributos não separa claramente os clusters. A variável *Mãos Abertas*, por exemplo, é praticamente a mesma nos três clusters. Alguns atributos têm pouca variação, como *Ereto*, sendo esse atributo importante, pois indica que o aluno não está com a cabeça baixa, o que é considerado um mau comportamento nas apresentações orais (CAVANAGH *et al.*, 2014). Outro atributo que separa os três clusters é *Mãos Baixas*. Três padrões claros podem ser vistos na Figura 10(a).

Na segunda apresentação, o comportamento geral permaneceu semelhante ao anterior. Por exemplo, temos que o atributo *Olhando o Público* corresponde praticamente ao mesmo comportamento. Nesta apresentação, também é possível identificar um grupo, o grupo C, com o atributo *Fala* com o valor mais alto (32% do tempo total da apresentação). Mas também pode ser observado que o atributo *Fala* possui valores mais próximos entre dois clusters do que na primeira apresentação. Também pode-se notar uma separação no atributo *Braços Cruzados*, com três grupos muito claros.

Nesta apresentação, o atributo *Mãos Baixas* também obteve uma variação entre três comportamentos.

Finalmente, na terceira apresentação deste ano, um novo comportamento, semelhante ao anterior, foi observado. Por outro lado, o atributo *Fala* apresenta dois grupos (A e B) com valores altos (cerca de 30 %), isso pode indicar que os apresentadores convergiram para um comportamento de fala mais ao longo do tempo (ano) em que eles estavam apresentando. Esses mesmos dois grupos ainda têm valores próximos ao atributo *Apontando* (aproximadamente 10 % para ambos). O atributo *Apontando* é relevante em apresentações orais com o auxílio de PowerPoint, pois apontar pode conectar as ideias do apresentador (KELLY *et al.*, 2008). Há também outro grupo que apresenta valores baixos para esses atributos, (*Fala e Apontando*).

5.3.2 Apresentações de 2018

Os atributos do ano de 2018, para a primeira apresentação, se comportaram de maneira um pouco diferente de 2017. Primeiro, pode-se supor que todos os grupos tenham o atributo *Fala* com valores inferiores a 10 % do total de apresentações. O atributo *Apontando*, como *Mãos Baixas*, possui uma separação entre os três clusters escolhidos. Para a postura *Braços Cruzados*, esse comportamento de separação nos três grupos também pode ser encontrado. Em geral, esta primeira apresentação não marca claramente três comportamentos diferentes. Aparentemente, os clusters são distribuídos pelos atributos de forma mais aleatória do que no ano de 2017.

Em 2018, os valores dos atributos entre os três clusters são distribuídos em uma escala. O grupo A teve valores mais altos, no grupo C, os valores foram um pouco menores e o grupo B teve valores menores que os dois anteriores. Assim, pode-se ver na Figura 10(d) que a primeira apresentação de 2018 teve seus clusters formando algo semelhante a um escala. Esse comportamento é diferente do que foi apresentado anteriormente (os clusters variaram de acordo com os centroides nas apresentações).

O comportamento inicial com alguma aleatoriedade nos dados e comportamentos na primeira apresentação, não ocorre na segunda apresentação. Ao analisar os valores dos dados, pode-se observar que os grupos são formados com uma "regra", ou comportamento, assim como foram formados em 2017. Esse comportamento é um cluster com maior valor em um atributo, outro cluster com um valor intermediário e um terceiro com valor inferior aos dois anteriores.

A última apresentação de 2018 se comporta de maneira diferente do resto do ano, como mostra a Figura 10(f). Primeiro, pode-se observar no gráfico que os valores do centroide diminuem em relação às apresentações anteriores. Outro ponto a ser observado é que os agrupamentos formados mudaram sua regra geral. Por exemplo, temos que o atributo *Olhando o Público* não possui mais dois clusters com valores altos e um com valor baixo. Nesta apresentação, dois clusters possuem valores baixos,

grupos B e C. Em geral, em 2018, não podemos observar o mesmo padrão que ocorre em 2017; a regra de formação de clusters não é a mesma, principalmente, na última apresentação do ano de 2018.

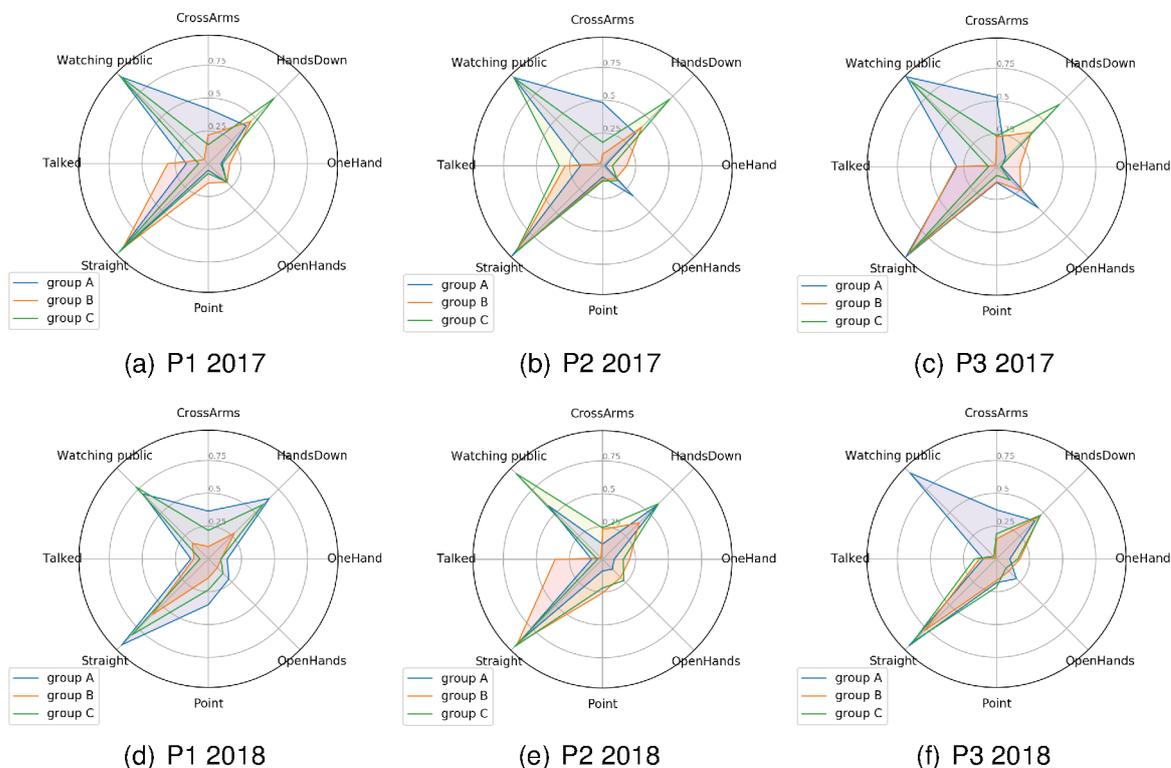


Figura 10 – Gráficos polares dos centroides de cada apresentação. Fonte: (ROQUE, F. *et al.*, 2019b)

5.3.3 Comparação Entre os Anos

É difícil apontar quais são os mesmos grupos (se houver) ao longo do ano, entre as apresentações. Nesse sentido, a comparação feita no presente trabalho será identificar o caminho percorrido ao longo do ano e, principalmente, uma comparação entre as últimas apresentações de cada ano. Com isso, podemos identificar algum padrão que foi formado durante o semestre. Os resultados da Tabela 7 mostram que, no final de cada semestre, o número de atributos estatisticamente diferentes cai de 6 para 1.

O ano de 2017 alcançou certa constância nos padrões. As apresentações variavam em valores de centroide, no entanto, o gráfico polar ajudou a encontrar três comportamentos nas três apresentações. Esse comportamento não pode ser observado para o ano de 2018, onde o mesmo comportamento (ou semelhança) entre apresentações não é evidente.

O que pode ser visto nos dois anos (2017 e 2018) é que há uma convergência de alguns atributos com maiores valores, na última apresentação. Por exemplo,

pode-se ver que o grupo A, em dois anos, tem um valor semelhante para a variável Braços Cruzados. Da mesma forma, o Grupo B possui atributos com valores semelhantes, como Olhando o Público e Braços Cruzados. O grupo C apresenta alguma variação entre as apresentações, mas é possível perceber uma semelhança entre as apresentações dos dois anos.

Depois de todos esses resultados, pode-se dizer que três grupos distintos foram encontrados nas apresentações de 2017, isso é fácil de observar na Figura 10. Para 2018, não foram observadas apresentações com os mesmos padrões de 2017. Para este ano, ficou mais difícil separar os clusters. Eles foram chamados de Grupo A, Grupo B e Grupo C. Essa escolha foi feita após uma análise das *Silhouettes* e suas médias. Também leva-se em consideração os valores dos centroides desses grupos ($k = 2$ retorna um valor melhor da média de *Silhouettes*), onde foi observada uma clara separação entre esses três aglomerados. Um número maior de clusters ($k = 5$ na primeira apresentação de 2018) também foi possível em casos específicos (Tabela 8). Mas para todas as outras apresentações, o valor médio da Silhouette foi melhor em $k = 3$. Isso pode ser chamado de semelhança entre os anos.

Semelhante em (SCHNEIDER; BLIKSTEIN, 2015), este estudo preocupou-se em encontrar padrões nas interações dos alunos em ambientes complexos de aprendizagem; neste caso, apresentações. Esse processo pode ajudar professores e educadores com uma maneira fácil e eficiente de analisar rapidamente as posturas corporais dos alunos.

Existem semelhanças gerais entre os dados coletados nos anos de 2017 e 2018. A Tabela 5 mostra o comportamento semelhante entre os dois anos para os valores médios. A Figura 7 mostra as semelhanças nos atributos avaliados. Especialmente, no início do semestre, observou-se que as apresentações variam. Isso pode ser explicado pelo número de observações encontradas na primeira apresentação de 2018 (59 observações), um número maior do que em 2017 (40 observações). O teste não paramétrico *Wilcoxon Rank Sum* ajuda a observar essa diferença. Na Tabela 7, o teste (comparação) que apresentou ocorrências de atributos mais significativas ocorreu entre as primeiras apresentações de 2017 e 2018 (6 atributos). Observa-se também que, ao final, com a comparação entre as duas últimas apresentações, o número desses atributos cai para 1. Isso indica que há uma maior igualdade estatística entre os alunos, ao final das apresentações.

Outro ponto do estudo a ser destacado é a importância do atributo *Fala*. Devido à configuração experimental, essa variável pode identificar mais facilmente determinados padrões ou mesmo um cluster em todo o conjunto de dados. Ainda assim, o atributo *Fala* pode destacar as outras características (atributos) dos alunos, uma vez que bons comunicadores também devem ter habilidades com gestos corporais (LUCAS, 2011).

5.4 SEQUENTIAL PATTERN MINING (SPM)

Foram avaliadas as apresentações individualmente, avaliando-se o suporte (quantidade de ocorrência da sequência na apresentação) de cada uma das sequências. Para cada apresentação de cada ano foi gerado um set de sequências mais frequentes. Os atributos utilizados no algoritmo de SPM são os mesmos do restante da experimentação, 3. O objetivo principal dessa etapa é identificar as principais sequências presentes nas apresentações e ainda se essas ocorrem em todas as apresentações ao longo do ano. Mostrar todos os padrões de sequências encontrados com o algoritmo se revela uma tarefa difícil. Assim, foram avaliados diferentes valores de suporte mínimo para os dois diferentes anos. Isso ocorre devido às características de cada base de dados. Ao avaliar o ano de 2017, um valor muito grande de sequências o que fez com que fosse necessário testar valores diferentes de suporte mínimo.

5.4.1 Aplicação do MINSUP 95% nas Apresentações de 2017

De modo geral o principal padrão que o algoritmo encontra, numa primeira avaliação, é o padrão contendo a sequência com um único atributo como a variável *Braços Cruzados*, por exemplo. Existem combinações em diferentes quantidades com esse atributo, sendo uma sequência de tempo que o estudante permanece nesse estado. Esses padrões de um único atributo não são tão relevantes para o estudo visto que o objetivo seria justamente entender o comportamento do estudante, não sendo interessante avaliar sequências onde não existem trocas de posturas (atributos).

Outro fato que pode ser observado ao avaliar as sequências encontradas pelo algoritmo é que muitas sequências terminam no atributo *Mãos Baixas* ou no atributo *Ereto*, como pode ser visto na Tabela 9. Ambos os atributos são formas comuns de estar durante uma apresentação: posição do corpo ereta e com as mão estendidas para baixo, ambas ao final das sequências.

Neste estudo foi decidido avaliar algumas sequências específicas do total que o algoritmo encontrou. Essas sequências avaliadas são consideradas mais relevantes e de certa forma representam o comportamento geral das outras sequências presentes no restante dos resultados.

Para a primeira apresentação, Tabela 9, algumas sequências interessantes apareceram como resultado. Por exemplo, a primeira sequência que aparece na tabela, com suporte 40. Nessa sequência, nota-se um comportamento mais passivo, alternando na postura de braços cruzados, mãos baixas, e mãos abertas (pode significar que o estudante está explicando algo) e por fim a posição ereta.

Ainda para a mesma apresentação, muitas sequências possuem o atributo *Braços Cruzados* presente em várias sequências. Isso pode acontecer por ser a primeira apresentação e os estudantes ainda não tinham grande habilidade de apresentação.

Ressalta-se que o atributo *Mãos Baixas* aparece em diversas sequências, sendo outro indicativo que o estudante passa muito tempo numa atitude passiva.

Outras sequências na Tabela 9 apresentam o atributo *Fala* com característica relevante. Na última sequência percebe-se que o atributo de fala é seguido pelo atributo *Uma Mão*, que indica que o estudante passa a apontar com uma mão, tentando explicar algo após um período de explanação verbal.

Tabela 9 – Sequências selecionadas da primeira apresentação de 2017.

Apresentação 1 (2017)	
Padrão	Sup
Braços Cruzados, Mãos Baixas, Mãos Abertas, Braços Cruzados, Ereto	40
Braços Cruzados, Mãos Baixas, Fala, Ereto	39
Braços Cruzados, Uma Mão, Mãos Abertas, Ereto	39
Braços Cruzados, Fala, Ereto	39
Mãos Baixas, Uma Mão, Fala, Ereto	39
Fala, Braços Cruzados, Mãos Baixas	39
Braços Cruzados, Fala, Uma Mão, Ereto	38
Mãos Baixas, Fala, Braços Cruzados, Mãos Abertas, Braços Cruzados, Mãos Baixas	38
Fala, Uma Mão, Mãos Abertas, Mãos Baixas, Ereto	38

Na mesma linha da primeira apresentação, a segunda apresentação apresenta comportamento parecido. Algumas variações dos atributos *Mãos Baixas* e *Mãos Abertas* em sequência com outros atributos de característica mais passiva. Esse resultado pode ser visto na Tabela 10.

O atributo *Uma Mão* tem destaque nessa apresentação, com suporte 21. Ele aparece em algumas sequências e particularmente na segunda sequência da Tabela 10 os estudantes tiveram um comportamento de estar com os atributos *Mão na Face*, *Mãos Abertas* e em algum momento ficar com a posição de *Uma Mão* apenas e retornar para *Mãos Abertas* novamente. Esse é um indicativo (indicado pelo comportamento de troca para uma mão) que os estudantes explicaram algo.

Ainda para a segunda apresentação, o atributo *Apontando* aparece com destaque. Com suporte 21, ele aparece acompanhado com o atributo *Uma Mão*. Já a variável *Fala* tem um comportamento interessante na quarta sequência da Tabela 10. Aparece no meio de outros atributos em momentos diferentes. Deduz-se que o estudante explicava por meio verbal e baixou as mãos, talvez explicando com as duas mãos (interação não verbal) algo, voltando para a explicação verbal.

Para a terceira apresentação, Tabela 11, as sequências mantêm o padrão encontrado nas apresentações anteriores. Destaque novamente para a grande presença do atributo *Braços Cruzados* nas sequências de modo geral. Como nas outras apresentações, existem muitos atributos apresentando sequências dele mesmo, que não foram expostos na tabela pelo motivo anterior elucidado.

Tabela 10 – Sequências selecionadas da segunda apresentação de 2017.

Apresentação 2 (2017)	
Padrão	Sup
Braços Cruzados, Uma Mão, Mãos Abertas, Ereto	21
Mão na Face, Mãos Baixas, Mãos Abertas, Mãos Baixas, Uma Mão, Mãos Abertas, Ereto	21
Mão na Cintura, Mãos Abertas, Ereto	21
Uma Mão, Mãos Abertas, Fala, Mãos Abertas, Fala, Ereto	21
Uma Mão, Apontando, Uma Mão, Ereto	21
Mãos Abertas, Uma Mão, Apontando, Mãos Abertas, Ereto	21

Destaca-se que para o último padrão da tabela da apresentação 3 de 2017, com suporte 22, os estudantes estavam apontando algo (os slides, possivelmente) passaram a ficar com as mãos abertas, passaram para a postura de uma mão apenas, terminando a sequência eretos (posição do corpo). De certa forma esse pode ser considerado um comportamento mais ativo dos estudantes, mais participativos em relação a ficar com os braços cruzados.

Tabela 11 – Sequências selecionadas da terceira apresentação de 2017.

Apresentação 3 (2017)	
Padrão	Sup
Braços Cruzados, Mãos Baixas, Mãos Abertas, Ereto	23
Braços Cruzados, Mãos Baixas, Mãos Abertas, Olhando o Público	23
Mãos Baixas, Uma Mão, Ereto	23
Mãos Baixas, Uma Mão, Olhando o Público	23
Uma Mão, Apontando, Mãos Abertas, Ereto	23
Mãos Baixas, Mãos Abertas, Apontando, Mãos Baixas, Ereto	22
Mãos Baixas, Uma Mão, Mãos Abertas, Olhando o Público	22
Apontando, Mãos Abertas, Uma Mão, Ereto	22

5.4.2 Aplicação do MINSUP 60% nas Apresentações de 2018

Diferentemente das apresentações de 2017, as apresentações de 2018 mostram sequências diferentes de maneira geral. Primeiro pela quantidade de itens nas sequências que é menor com relação ao ano de 2017. Isso não significa que não existam sequências grandes nas apresentações de 2018. Assim como nas apresentações de 2017, existem sequências com um único item repetidas vezes. Em várias sequências encontradas com 8 vezes o mesmo item.

De modo geral, os atributos que aparecem com mais frequência nas apresentações de 2017 também aparecem nas apresentações de 2018. Exemplos são os atributos: *Mãos Baixas*, *Uma Mão*, *Ereto* e *Mãos Abertas*. Como pode ser visto, essa diferença entre as apresentações de 2018 e as apresentações de 2017 pode ser expli-

cada pela diferença estatística observada na Tabela 7. Na tabela, principalmente entre as duas primeiras apresentações de cada ano, existem 6 atributos estatisticamente diferentes.

A primeira sequência da Tabela 12, com suporte 53, apresenta dois atributos considerados de atitudes passivas por parte dos estudantes, *Mãos Baixas* e *Ereto*. Pelo valor alto do suporte, essa sequência aparece em muitas apresentações.

Tabela 12 – Sequências selecionadas da primeira apresentação de 2018.

Apresentação 1 (2018)	
Padrão	Sup
Mãos Baixas, Ereto	53
Olhando o Público, Mãos Abertas	47
Uma Mão, Uma Mão, Mãos Abertas	45
Apontando, Mãos Baixas, Mãos Baixas	43
Mãos Baixas, Mãos Baixas, Ereto	42
Mãos Baixas, Ereto, Olhando o Público	41
Uma Mão, Uma Mão, Braços Cruzados, Braços Cruzados	41

A seguir, ao avaliar a segunda apresentação de 2018 percebe-se que ela mantém a característica da primeira apresentação em não possuir sequências de itens com variação de atributos grande. Esse resultado pode ser visto na Tabela 13.

Com suporte 37, a primeira sequência da tabela da segunda apresentação pode ser considerada de certa forma interessante. Os estudantes deixam de estar com a posição de uma mão, para colocar a Olhando o Público. A princípio isso pode ser uma transição, ou movimento, para colocar a mão no rosto. O software interpreta como uma troca de posição e ação e grava na base de dados.

Ainda nessa apresentação, a terceira sequência da Tabela 13 apresenta um padrão que pode ser observado nas apresentações de 2017. Sendo uma troca entre a posição de braços de uma mão para a posição apontando, com suporte de 36.

A última sequência da tabela também representa um padrão interessante e que foi observado nas apresentações de 2017. A sequência de dois intervalos de fala com um de uma mão (podendo indicar explicação). Essa sequência também se destaca pelo suporte de 30.

Tabela 13 – Sequências selecionadas da segunda apresentação de 2018.

Apresentação 2 (2018)	
Padrão	Sup
Uma Mão, Mão na Face	37
Uma Mão, Mãos Abertas	37
Uma Mão, Apontando	36
Ereto, Mão na Face	31
Fala, Fala, Uma Mão	30

Por fim, a última apresentação de 2018 com as sequências apresentadas na Tabela 14, alguns padrões com o atributo *Braços Cruzados* voltam a ter destaque. Este pode ser considerado padrão similar com as apresentações de 2017, ressaltando que ao final, pelos dados estatísticos da Tabela 7, as apresentações devem ser mais similares, já que o número de atributos diferentes é apenas 1.

Tabela 14 – Sequências selecionadas da terceira apresentação de 2018.

Apresentação 3 (2018)	
Padrão	Sup
Braços Cruzados, Mãos Baixas	25
Mão na Face, Mãos Baixas	25
Mãos Abertas, Uma Mão	25
Mãos Abertas, Mãos Abertas, Mãos Baixas	25
Mão na Face, Mão na Face, Mãos Baixas	24
Braços Cruzados, Mãos Baixas	23
Braços Cruzados, Apontando	23

5.5 DISCUSSÃO

A proposta de um Framework Conceitual para análise de posturas corporais pode guiar trabalhos futuros de outros autores (WORSLEY, 2018). De um lado, ele pode fornecer uma compreensão geral dos dados que estão sendo utilizados. Ao mesmo tempo que cada processo que é feito nos dados retorna um resultado diferente, agregando conhecimento e entendimento do problema estudado.

A análise descritiva dos dados pode tornar mais claro o problema em relação aos atributos avaliados. Como pode ser visto na Tabela 5, a média aritmética simples pode deixar mais evidente que o atributo *Ereto* aparece com muita frequência. Da mesma maneira, a análise temporal, mostrada na Figura 7 apresenta uma relação entre os anos. Os gráficos de tendência mostram certo padrão entre as apresentações no mesmo momento.

A análise com a estatística inferencial pode encontrar informações mais relevantes para o estudo. Como apresentado na Tabela 7, essa comparação entre os anos e apresentações com o teste *Wilcoxon rank sum* pôde mostrar, por exemplo que comparando cada ano em cada apresentação em mesmo momento, poucos atributos são diferentes estatisticamente. Contudo, ao comparar os dois anos, para a primeira apresentação metade dos atributos são estatisticamente diferentes. Ainda, pode-se observar que na última apresentação tem-se apenas 1 atributo estatisticamente diferente, representando um convergência, indicando assim que ao final, na última apresentação, os estudantes estão apresentando de maneira mais igual entre os dois anos.

A utilização do algoritmo de agrupamento *k-means* possibilita identificar várias características dos dados. O algoritmo é uma ferramenta importante que segmenta

dados com base em um modelo matemático. Pode-se, com o uso do *k-means*, identificar os grupos de comportamentos, com base nos atributos, gerados pelos estudantes ao apresentar seus trabalhos. A análise dos centroides, Figura 10, possibilita essa compreensão geral desses comportamentos. Contudo, sozinho o algoritmo *k-means* se torna pouco efetivo, visto que escolher o número certo para *k* pode ser uma tarefa difícil. As *Silhouettes*, Figuras 8 e 9, ajudam tanto de forma visual como de forma numérica (Tabela 8), a encontrar o valor que mais se adéqua ao problema.

As sequências, usando o algoritmo *PrefixSpan*, evidenciam certos comportamentos que sem o recurso do algoritmo, são de difícil visualização. Por exemplo, na apresentação 3 de 2017, Tabela 11, a primeira sequência sendo {*BraçosCruzados*, *MãosBaixas*, *MãosAbertas*, *Ereto*}, o aluno está em uma posição de braços menos ativa e durante toda essa sequência. O que pode ser visto anteriormente nos resultados dos centroides (identificação de diferentes tipos de comportamentos). Esse resultado também pode ser relacionado com os resultados de estatística descritiva, os atributos que mais ocorrem, por exemplo, *Ereto*, também ocorrem com frequência nas sequências.

A análise das apresentações com dados multimodais possibilita uma melhor compreensão do processo de aprendizagem. É difícil estimar um estado mental complexo baseado somente em uma característica de comportamento (CHAN *et al.*, 2020). O uso do kinect com seus sensores possibilitou por exemplo, identificar padrões de apresentação diferentes entre o grupo total avaliado dos dados, identificados pelo *k-means*. No mesmo sentido, Chan *et al.* (2020) sugerem que o MMLA habilita uma extração de detalhes que seria muito difícil de perceber por olhos humanos. Características que anteriormente não seriam possíveis ou não seria preciso de observar, a exemplo das sequências mais frequentes dos estudantes.

Para Spikol e Cukurova (2019), como objetivo atual, o MMLA precisa ainda ser integrado em ambientes do mundo real, onde ambientes de educação são ecologicamente válidos. Nesse sentido a integração do Framework apresentado com um LMS padrão poderia fornecer informações valiosas visando tanto a avaliação individual do aluno, como de uma turma ou turmas em diferentes anos. Ainda, a análise de sequências em um ambiente real, com mais informações como nota e impressões do professor, poderia diferenciar uma boa apresentação de uma ruim, visto que com esse resultado pode-se observar as sequências mais frequentes nas apresentações e rastrear quais estudantes fizeram quais sequências.

O modelo de avaliação que o Framework Conceitual cria, ajuda a revelar traços do comportamento do estudante a medida que ele vai apresentando. Também é possível identificar padrões dentro das apresentações (sem a comparação entre os anos no estudo de caso apresentado), com a avaliação dos centroides. Ao final, passando por todas as etapas, uma base de dados contendo os intervalos de tempo que o estudante

passa em cada postura, deve possuir como resultado um aluno com características de apresentação bem definidas. Ainda, se existirem dois ou mais anos para avaliação, a relação (semelhança entre atributos) entre os anos poderá evidenciar essa relação de maneira estatística. Dessa forma, essa comparação e identificação do aluno, pode ser usada para atribuir ganho ou perda de aprendizagem, visto que isso pode estar inserido junto com a avaliação do professor ou professores.

6 CONCLUSÃO

Um Framework Conceitual que usa MMLA pode auxiliar na avaliação de ambientes complexos de aprendizado usando dados de várias fontes. No presente estudo, constata-se que durante as apresentações avaliadas pelo teste estatístico inferencial, comparado entre os anos de 2017 e 2018, a quantidade de atributos que possuem diferença estatística de diminuiu de 6 para 1. Isso indica que, no final do semestre, os alunos apresentam de maneira mais semelhante entre si. Os resultados também indicam que para $k = 3$, existem 3 grupos diferentes nas apresentações. No ano de 2017, pode-se ver todos os três comportamentos mais claramente do que em 2018, onde a regra de grupos não é traduzida muito nítida entre os atributos, para todas as apresentações.

De maneira geral, algumas limitações foram encontradas ao realizar o trabalho. Sobretudo, na tentativa de comparar os padrões criados pelos centroides quando $k = 3$. Ao avaliar o ano de 2017 era possível observar certo padrão de formação dos grupos. Ainda, pode-se avaliar o comportamento ao longo das 3 apresentações do ano de 2017, mas isso não foi possível para o ano de 2018. A formação dos grupos, por conta das características próprias dos dados de 2018, não apresenta uma maneira clara de formação. Isso dificultou a comparação entre os anos e não permitiu uma evolução nesse sentido.

Outro ponto identificado como uma limitação, natural do aspecto dos dados, é a falta de outros tipos de informações que não a de posturas corporais e fala. Seria importante para o estudo, por exemplo, as notas ou impressões do professor estarem presentes. Essas informações poderiam ser usadas como forma de fazer outros apontamentos. Nas sequências uma visão de qualidade de apresentação poderia ser dado com mais precisão. Ainda poderia ser criado um modelo de classificação ou regressão se existissem notas.

Como resultado geral da pesquisa, também podemos observar que o método visual de *Silhouettes*, juntamente com a avaliação de centroides, ajuda a avaliar esse tipo de problema. Com essas técnicas, foi possível observar a formação de grupos com os centroides. Ainda, com as *Silhouettes*, foi possível determinar o melhor valor de k para o problema, de uma maneira geral. Note-se que os sensores do Microsoft Kinect conseguiram criar um banco de dados que permitia a separação entre grupos (clusters).

O teste inferencial entre os atributos também mostrou que as comparações nas apresentações do mesmo ano apresentaram poucos atributos com diferenças estatísticas. Pode-se tomar como exemplo que a comparação entre a segunda apresentação de 2018 e a terceira apresentação de 2018 que não resultou em nenhum atributo estatisticamente diferente.

Existem dois principais resultados importantes da capacidade de identificar os padrões corporais do aluno durante as apresentações e analisar seus dados: um relacionado à avaliação do aluno e outro relacionado à criação de uma ferramenta para analisar o comportamento de diferentes grupos. O primeiro resultado é que a captura de apresentação e a detecção de padrões corporais permitem a auto-avaliação do aluno. Por exemplo, se um aluno estiver falando a maior parte do tempo disponível e não estiver usando a comunicação não verbal de maneira eficiente, um relatório automático fornecido como *feedback* pode ajudá-lo a avaliar melhor sua apresentação. No caso de uma avaliação regular em que as habilidades de apresentação sejam importantes, além do *feedback* qualitativo do professor, ele pode usar as informações adquiridas para fornecer uma profundidade mais quantitativa para articular suas recomendações para melhorar o desempenho do aluno.

Por outro lado, a capacidade de analisar dados estatísticos de grupos durante apresentações pode ser usada como uma ferramenta para comparar perfis de alunos em diferentes campos ou gerações. As informações podem ser usadas para definir estratégias educacionais adequadas. Indo além das possibilidades atuais do estudo, se um modelo para avaliar as habilidades de apresentação for criado ou adaptado pelos psicólogos para usar esse tipo de entrada, isso poderá se tornar uma ferramenta poderosa para avaliar como fatores internos e externos podem determinar a qualidade ou o estilo da apresentação.

6.1 TRABALHOS FUTUROS

Para trabalhos futuros, novas análises devem ser realizadas em diferentes grupos de estudantes. Esses novos alunos viriam de diferentes áreas de especialização, como engenharia, ciências da saúde e ciências humanas. Essa distinção pode mostrar comportamentos diferentes entre as áreas do conhecimento ou revelar comportamentos semelhantes entre eles. Aliadas a isso, as técnicas de visualização da informação poderiam ser exploradas para ajudar na detecção de padrões obtidos dos alunos provenientes das diferentes áreas do conhecimento. Como forma de devolver *feedback* imediato para os alunos, os tipos de perfis podem ser integrados ao software para classificação automática de apresentações em tempo real. Também é esperado testar outras técnicas de classificação para melhorar o desempenho de modelos.

A aquisição de outros tipos de dados, como notas, ou aspectos dos estudantes que possam caracterizar melhor a apresentação do estudantes podem tornar as análises mais abrangentes. Como foi mencionado, podem ser usados classificadores, bem como a relação entre as sequências e o comportamento pode ser melhor avaliado, visto que neste estudo, essa comparação ficou limitada aos dados que existiam. Métodos neurais também podem ser usados para comparação entre os resultados.

REFERÊNCIAS

ARTHUR, David; VASSILVITSKII, Sergei. **k-means++: The advantages of careful seeding**. [S.l.], 2006.

BECK, Joseph E; SISON, June. Using knowledge tracing in a noisy environment to measure student reading proficiencies. **International Journal of Artificial Intelligence in Education**, IOS Press, v. 16, n. 2, p. 129–143, 2006.

BEYAN, Cigdem; CARISSIMI, Nicolò; CAPOZZI, Francesca; VASCON, Sebastiano; BUSTREO, Matteo; PIERRO, Antonio; BECCHIO, Cristina; MURINO, Vittorio. Detecting emergent leader in a meeting environment using nonverbal visual features only. *In*: PROCEEDINGS of the 18th ACM International Conference on Multimodal Interaction. [S.l.: s.n.], 2016. P. 317–324.

BIDWELL, Jon; FUCHS, Henr. **Classroom Analytics: Measuring Student Engagement with Automated Gaze Tracking**. Chapel Hill, NC, 2011.

BLIKSTEIN, Paulo. Multimodal learning analytics. *In*: ACM. PROCEEDINGS of the third international conference on learning analytics and knowledge. [S.l.: s.n.], 2013. P. 102–106.

BLIKSTEIN, Paulo; WORSLEY, Marcelo. Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. **Journal of Learning Analytics**, v. 3, n. 2, p. 220–238, 2016.

BURGOON, Judee K; HOOBLER, Gregory D. Nonverbal signals. **Handbook of interpersonal communication**, v. 2, p. 229–285, 1994.

CAVANAGH, Michael; BOWER, Matt; MOLONEY, Robyn; SWELLER, Naomi. The effect over time of a video-based reflection system on preservice teachers' oral presentations. **Australian Journal of Teacher Education**, v. 39, n. 6, p. 1, 2014.

CHAN, Man Ching Esther; OCHOA, Xavier; CLARKE, David. Multimodal learning analytics in a laboratory classroom. *In*: MACHINE Learning Paradigms. [S.l.]: Springer, 2020. P. 131–156.

CHELLA, A.; DINDO, H.; INFANTINO, I. A System for Simultaneous People Tracking and Posture Recognition in the context of Human-Computer Interaction. *In*:

EUROCON 2005 - The International Conference on "Computer as a Tool". [S.l.: s.n.], nov. 2005. P. 991–994. DOI: 10.1109/EURCON.2005.1630114.

CHEN, Lei; FENG, Gary; JOE, Jilliam; LEONG, Chee Wee; KITCHEN, Christopher; LEE, Chong Min. Towards automated assessment of public speaking skills using multimodal cues. *In: PROCEEDINGS of the 16th International Conference on Multimodal Interaction*. [S.l.: s.n.], 2014. P. 200–203.

CHEN, Lei; LEONG, Chee Wee; FENG, Gary; LEE, Chong Min; SOMASUNDARAN, Swapna. Utilizing multimodal cues to automatically evaluate public speaking performance. *In: IEEE. AFFECTIVE Computing and Intelligent Interaction (ACII), 2015 International Conference on*. [S.l.: s.n.], 2015. P. 394–400.

COHEN, I.; LI, H. Inference of human postures by classification of 3D human body shape. *In: 2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443)*. [S.l.: s.n.], out. 2003. P. 74–81. DOI: 10.1109/AMFG.2003.1240827.

CUKUROVA, Mutlu; LUCKIN, Rose; MILLÁN, Eva; MAVRIKIS, Manolis. The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. **Computers & Education**, Elsevier, v. 116, p. 93–109, 2018.

D'MELLO, Sidney K.; CRAIG, Scotty D.; WITHERSPOON, Amy; MCDANIEL, Bethany; GRAESSER, Arthur. Automatic detection of learner's affect from conversational cues. **User Modeling and User-Adapted Interaction**, v. 18, n. 1, p. 45–80, fev. 2008. ISSN 1573-1391. DOI: 10.1007/s11257-007-9037-6. Disponível em: <https://doi.org/10.1007/s11257-007-9037-6>.

DE KONING, Björn B; TABBERS, Huib K; RIKERS, Remy MJP; PAAS, Fred. Attention guidance in learning from a complex animation: Seeing is understanding? **Learning and instruction**, Elsevier, v. 20, n. 2, p. 111–122, 2010.

DUDA, Richard O; HART, Peter E; STORK, David G. **Pattern classification**. [S.l.]: John Wiley & Sons, 2012.

DUDA, Richard O; HART, Peter E; STORK, David G. **Pattern classification and scene analysis 2nd ed.** ed: **Wiley Interscience**, 1995.

DUIN, Robert PW; TAX, DMJ. Statistical pattern recognition. *In: HANDBOOK of Pattern Recognition and Computer Vision*. [S.l.]: World Scientific, 2005. P. 3–24.

ECHEVERRIA, Vanessa; AVENDAÑO, Allan; CHILUIZA, Katherine; VÁSQUEZ, Aníbal; OCHOA, Xavier. Presentation skills estimation based on video and kinect data analysis. *In: ACM. PROCEEDINGS of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge. [S.l.: s.n.], 2014. P. 53–60.*

ECHEVERRIA, Vanessa; MARTINEZ-MALDONADO, Roberto; GRANDA, Roger; CHILUIZA, Katherine; CONATI, Cristina; SHUM, Simon Buckingham. Driving data storytelling from learning design. *In: PROCEEDINGS of the 8th international conference on learning analytics and knowledge. [S.l.: s.n.], 2018. P. 131–140.*

FACELI, Katti; LORENA, Ana Carolina; GAMA, João; CARVALHO, André Carlos Ponce de Leon *et al.* Inteligência Artificial: Uma abordagem de aprendizado de máquina, 2011.

FOURNIER-VIGER, Philippe; LIN, Jerry Chun-Wei; GOMARIZ, Antonio; GUENICHE, Ted; SOLTANI, Azadeh; DENG, Zhihong; LAM, Hoang Thanh. The SPMF open-source data mining library version 2. *In: SPRINGER. JOINT European conference on machine learning and knowledge discovery in databases. [S.l.: s.n.], 2016. P. 36–40.*

FOURNIER-VIGER, Philippe; LIN, Jerry Chun-Wei; KIRAN, Rage Uday; KOH, Yun Sing; THOMAS, Rincy. A survey of sequential pattern mining. **Data Science and Pattern Recognition**, v. 1, n. 1, p. 54–77, 2017.

FREUND, Yoav; SCHAPIRE, Robert E. **A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting.** [S.l.: s.n.], 1995.

FRISCHEN, Alexandra; BAYLISS, Andrew P; TIPPER, Steven P. Gaze cueing of attention: visual attention, social cognition, and individual differences. **Psychological bulletin**, American Psychological Association, v. 133, n. 4, p. 694, 2007.

GAN, Tian; WONG, Yongkang; MANDAL, Bappaditya; CHANDRASEKHAR, Vijay; KANKANHALLI, Mohan S. Multi-sensor Self-Quantification of Presentations. *In: PROCEEDINGS of the 23rd ACM International Conference on Multimedia. Brisbane, Australia: ACM, 2015. (MM '15), p. 601–610. DOI: 10.1145/2733373.2806252.* Disponível em: <http://doi.acm.org/10.1145/2733373.2806252>.

GIBBONS, Jean Dickinson; CHAKRABORTI, Subhabrata. Nonparametric statistical inference. *In: INTERNATIONAL encyclopedia of statistical science*. [S.l.]: Springer, 2011. P. 977–979.

GOMES, July; YASSINE, Mohamed; WORSLEY, Marcelo; BLIKSTEIN, Paulo. Analysing engineering expertise of high school students using eye tracking and multimodal learning analytics. *In: EDUCATIONAL Data Mining 2013*. [S.l.: s.n.], 2013.

HALL, Edward T. *et al.* Proxemics [and Comments and Replies]. **Current Anthropology**, [University of Chicago Press, Wenner-Gren Foundation for Anthropological Research], v. 9, n. 2/3, p. 83–108. ISSN 00113204, 15375382.

HART, John; GRATCH, Jonathan; MARSELLA, Stacy. How virtual reality training can win friends and influence people. **Human Factors in Defence**. Ashgate, p. 235–249, 2013.

HOWISON, Mark; TRNINIC, Dragan; REINHOLZ, Daniel; ABRAHAMSON, Dor. The Mathematical Imagery Trainer: from embodied interaction to conceptual learning. *In: PROCEEDINGS of the SIGCHI Conference on Human Factors in Computing Systems*. [S.l.: s.n.], 2011. P. 1989–1998.

HYMES, Dell. On communicative competence. **sociolinguistics**, Harmondsworth, v. 269293, p. 269–293, 1972.

IMENDA, Sitwala. Is there a conceptual difference between theoretical and conceptual frameworks? **Journal of Social Sciences**, Taylor & Francis, v. 38, n. 2, p. 185–195, 2014.

JABAREEN, Yosef. Building a conceptual framework: philosophy, definitions, and procedure. **International journal of qualitative methods**, SAGE Publications Sage CA: Los Angeles, CA, v. 8, n. 4, p. 49–62, 2009.

JERMANN, Patrick; MULLINS, Dejana; NÜSSLI, Marc-Antoine; DILLENBOURG, Pierre. Collaborative gaze footprints: Correlates of interaction quality. International Society of the Learning Sciences, 2011.

JEWITT, Carey. **Technology, literacy and learning: A multimodal approach**. [S.l.]: Psychology Press, 2006.

JOHNSON, Richard A. Statistical Inference. *In: International Encyclopedia of Statistical Science*. Edição: Miodrag Lovric. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. P. 1418–1420. ISBN 978-3-642-04898-2.

KELLY, Spencer D; MANNING, Sarah M; RODAK, Sabrina. Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. **Language and Linguistics Compass**, Wiley Online Library, v. 2, n. 4, p. 569–588, 2008.

KRESS, Gunther; VAN LEEUWEN, Theo. Multimodal discourse. **The Modes and Media of Contemporary Communication.(Cappelen, London 2001)**, 2001.

LE, Thi-Lan; NGUYEN, Minh-Quoc *et al.* Human posture recognition using human skeleton provided by Kinect. *In: IEEE. COMPUTING, Management and Telecommunications (ComManTel), 2013 International Conference on*. [S.l.: s.n.], 2013. P. 340–345.

LEONG, Chee Wee; CHEN, Lei; FENG, Gary; LEE, Chong Min; MULHOLLAND, Matthew. Utilizing depth sensors for analyzing multimodal presentations: Hardware, software and toolkits. *In: ACM. PROCEEDINGS of the 2015 ACM on International Conference on Multimodal Interaction*. [S.l.: s.n.], 2015. P. 547–556.

LIEHR, Patricia; SMITH, Mary Jane. Middle range theory: Spinning research and practice to create knowledge for the new millennium. **Advances in Nursing Science**, LWW, v. 21, n. 4, p. 81–91, 1999.

LILLIEFORS, Hubert W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. **Journal of the American statistical Association**, Taylor & Francis Group, v. 62, n. 318, p. 399–402, 1967.

LLOYD, Stuart. Least squares quantization in PCM. **IEEE transactions on information theory**, IEEE, v. 28, n. 2, p. 129–137, 1982.

LUCAS, Stephen. **The Art of Public Speaking, 11th Edition**. [S.l.]: McGraw-Hill Education, out. 2011. ISBN 007769290X.

LUZARDO, Gonzalo; GUAMÁN, Bruno; CHILUIZA, Katherine; CASTELLS, Jaime; OCHOA, Xavier. Estimation of Presentations Skills Based on Slides and Audio

Features. *In: PROCEEDINGS of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge*. Istanbul, Turkey: ACM, 2014. (MLA '14), p. 37–44. DOI: 10.1145/2666633.2666639. Disponível em: <http://doi.acm.org/10.1145/2666633.2666639>.

MARTINEZ-MALDONADO, Roberto; ECHEVERRIA, Vanessa; SANTOS, Olga C; SANTOS, Augusto Dias Pereira Dos; YACEF, Kalina. Physical learning analytics: A multimodal perspective. *In: PROCEEDINGS of the 8th International Conference on Learning Analytics and Knowledge*. [S.l.: s.n.], 2018. P. 375–379.

MASON, Lucia; PLUCHINO, Patrik; TORNATORA, Maria Caterina; ARIASI, Nicola. An eye-tracking study of learning from science text with concrete and abstract illustrations. **The Journal of Experimental Education**, Taylor & Francis, v. 81, n. 3, p. 356–384, 2013.

MEHRABIAN, Albert. **Nonverbal Communication**. [S.l.]: Routledge, jul. 2017. DOI: 10.4324/9781351308724. Disponível em: <https://doi.org/10.4324/9781351308724>.

MO, Hao-Cheng; LEOU, Jin-Jang; LIN, Cheng-Shian. Human Behavior Analysis Using Multiple 2D Features and Multicategory Support Vector Machine. *In: MVA2009 IAPR Conference on Machine Vision Applications, Proceedings*. [S.l.: s.n.], 2009.

MOGHADDAM, Zia; PICCARDI, Massimo. Human Action Recognition with MPEG-7 Descriptors and Architectures. *In: PROCEEDINGS of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams*. Firenze, Italy: ACM, 2010. (ARTEMIS '10), p. 63–68. DOI: 10.1145/1877868.1877885. Disponível em: <http://doi.acm.org/10.1145/1877868.1877885>.

MUNOZ, Roberto; VILLARROEL, Rodolfo; BARCELOS, Thiago S.; SOUZA, Alexandra; MERINO, Erick; GUIÑEZ, Rodolfo; SILVA, Leandro A. Development of a Software that Supports Multimodal Learning Analytics: A Case Study on Oral Presentations. **Journal of Universal Computer Science**, v. 24, n. 2, p. 149–170, 28 fev. 2018.

MUNTEANU, Cosmin; PENN, Gerald; ZHU, Xiaodan. Improving automatic speech recognition for lectures through transformation-based rules learned from minimal data. *In: PROCEEDINGS of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*. [S.l.: s.n.], 2009. P. 764–772.

NOEL, Rene; RIQUELME, Fabián; MAC LEAN, Roberto; MERINO, Erick; CECHINEL, Cristian; BARCELOS, Thiago S; VILLARROEL, Rodolfo; MUNOZ, Roberto. Exploring Collaborative Writing of User Stories With Multimodal Learning Analytics: A Case Study on a Software Engineering Course. **IEEE Access**, IEEE, v. 6, p. 67783–67798, 2018.

OCHOA, Xavier. Multimodal Learning Analytics. *In*: LANG, Charles; SIEMENS, George; WISE, Alyssa Friend; GAŠEVIC, Dragan (Ed.). **The Handbook of Learning Analytics**. 1. ed. Alberta, Canada: Society for Learning Analytics Research (SoLAR), 2017. P. 129–141. ISBN 978-0-9952408-0-3. Disponível em: <http://solaresearch.org/hla-17/hla17-chapter1>.

OCHOA, Xavier; CHILUIZA, Katherine; MÉNDEZ, Gonzalo; LUZARDO, Gonzalo; GUAMÁN, Bruno; CASTELLS, James. Expertise Estimation Based on Simple Multimodal Features. *In*: PROCEEDINGS of the 15th ACM on International Conference on Multimodal Interaction. Sydney, Australia: ACM, 2013. (ICMI '13), p. 583–590. DOI: 10.1145/2522848.2533789. Disponível em: <http://doi.acm.org/10.1145/2522848.2533789>.

OCHOA, Xavier; DOMINGUEZ, Federico; GUAMÁN, Bruno; MAYA, Ricardo; FALCONES, Gabriel; CASTELLS, Jaime. The RAP System: Automatic Feedback of Oral Presentation Skills Using Multimodal Analysis and Low-cost Sensors. *In*: PROCEEDINGS of the 8th International Conference on Learning Analytics and Knowledge. Sydney, New South Wales, Australia: ACM, 2018. (LAK '18), p. 360–364. DOI: 10.1145/3170358.3170406. Disponível em: <http://doi.acm.org/10.1145/3170358.3170406>.

OCHOA, Xavier; WORSLEY, Marcelo. Augmenting Learning Analytics with Multimodal Sensory Data. **Journal of Learning Analytics**, Society for Learning Analytics Research, v. 3, n. 2, p. 213–219, set. 2016. DOI: 10.18608/jla.2016.32.10. Disponível em: <https://doi.org/10.18608/jla.2016.32.10>.

OCHOA, Xavier; WORSLEY, Marcelo; WEIBEL, Nadir; OVIATT, Sharon. Multimodal learning analytics data challenges. *In*: PROCEEDINGS of the Sixth International Conference on Learning Analytics & Knowledge - LAK16. [S.l.]: ACM Press, 2016. DOI: 10.1145/2883851.2883913. Disponível em: <https://doi.org/10.1145/2883851.2883913>.

OVIATT, Sharon; COHEN, Adrienne. Written and multimodal representations as predictors of expertise and problem-solving success in mathematics. *In: ACM. PROCEEDINGS of the 15th ACM on International conference on multimodal interaction*. [S.l.: s.n.], 2013. P. 599–606.

RACA, Mirko; DILLENBOURG, Pierre. System for Assessing Classroom Attention. *In: PROCEEDINGS of the Third International Conference on Learning Analytics and Knowledge*. Leuven, Belgium: ACM, 2013. (LAK '13), p. 265–269. DOI: 10.1145/2460296.2460351. Disponível em: <http://doi.acm.org/10.1145/2460296.2460351>.

RACA, Mirko; TORMEY, Roland; DILLENBOURG, Pierre. Sleepers' Lag - Study on Motion and Attention. *In: PROCEEDINGS of the Fourth International Conference on Learning Analytics And Knowledge*. Indianapolis, Indiana, USA: ACM, 2014. (LAK '14), p. 36–43. DOI: 10.1145/2567574.2567581.

RASIPURAM, Sowmya; JAYAGOPI, Dinesh Babu. Automatic multimodal assessment of soft skills in social interactions: a review. **Multimedia Tools and Applications**, Springer, p. 1–24, 2020.

REILLY, Joseph; RAVENELL, Milan; SCHNEIDER, Bertrand. Exploring Collaboration Using Motion Sensors and Multi-Modal Learning Analytics. *In: PROCEEDINGS of the 11th International Conference on Educational Data Mining*. [S.l.: s.n.], 2018.

RICKHEIT, Gert; STROHNER, Hans. **Handbook of communication competence**. [S.l.]: Walter de Gruyter, 2008. v. 1.

RIQUELME, Fabian; MUNOZ, Roberto; LEAN, Roberto Mac; VILLARROEL, Rodolfo; BARCELOS, Thiago S.; ALBUQUERQUE, Victor Hugo C. de. Using multimodal learning analytics to study collaboration on discussion groups. **Universal Access in the Information Society**, Springer Science e Business Media LLC, jul. 2019. DOI: 10.1007/s10209-019-00683-w. Disponível em: <https://doi.org/10.1007/s10209-019-00683-w>.

ROCCO, Tonette S; PLAKHOTNIK, Maria S. Literature reviews, conceptual frameworks, and theoretical frameworks: Terms, functions, and distinctions. **Human Resource Development Review**, SAGE Publications Sage CA: Los Angeles, CA, v. 8, n. 1, p. 120–130, 2009.

ROQUE, Felipe Vieira; CECHINEL, Cristian; MERINO, Erick; VILLARROEL, Rodolfo; LEMOS, Robson; MUNOZ, Roberto. Using Multimodal Data to Find Patterns in Student Presentations. *In: IEEE. 2018 XIII Latin American Conference on Learning Technologies (LACLO). [S.l.: s.n.], 2018. P. 256–263.*

ROQUE, Felipe; CECHINEL, Cristian; MUÑOZ, Roberto; LEMOS, Robson; MERINO, Erick; ACEVEDO, Rodolfo Villarroel. Evolução das posturas corporais de estudantes em apresentações de seminários ao longo do semestre: uma análise utilizando dados multimodais e técnicas de clusterização. *In: 1. BRAZILIAN Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). [S.l.: s.n.], 2018. P. 1483.*

ROQUE, Felipe; CECHINEL, Cristian; MUÑOZ, Roberto; LEMOS, Robson; WEBER, Tiago Oliveira. Encontrando os padrões sequenciais em apresentações orais de estudantes utilizando Sequential Pattern Mining. *In: 1. BRAZILIAN Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). [S.l.: s.n.], 2019a. P. 1896.*

ROQUE, Felipe; CECHINEL, Cristian; WEBER, Tiago O; LEMOS, Robson; VILLARROEL, Rodolfo; MIRANDA, Diego; MUNOZ, Roberto. Using Depth Cameras to Detect Patterns in Oral Presentations: A Case Study Comparing Two Generations of Computer Engineering Students. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 19, n. 16, p. 3493, 2019b.

ROUSSEEUW, Peter J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of computational and applied mathematics**, Elsevier, v. 20, p. 53–65, 1987.

RUSSELL, Stuart J; NORVIG, Peter. **Artificial intelligence: a modern approach**. [S.l.]: Malaysia; Pearson Education Limited, 2016.

SARAF, Pratik; SEDAMKAR, RR; RATHI, Sheetal. PrefixSpan Algorithm for Finding Sequential Pattern with Various Constraints. **International Journal of Applied Information Systems (IJ AIS)**, Citeseer, p. 2249–0868, 2015.

SCHERER, Stefan; WORSLEY, Marcelo; MORENCY, Louis Philippe. 1st international workshop on multimodal learning analytics. English (US). *In: ICMI'12 - Proceedings of the ACM International Conference on Multimodal Interaction. [S.l.: s.n.], dez. 2012. DOI: 10.1145/2388676.2388803.*

SCHLÖMER, Thomas; POPPINGA, Benjamin; HENZE, Niels; BOLL, Susanne. Gesture recognition with a Wii controller. *In: PROCEEDINGS of the 2nd international conference on Tangible and embedded interaction*. [S.l.: s.n.], 2008. P. 11–14.

SCHNEIDER, Bertrand; BLIKSTEIN, Paulo. Unraveling students' interaction around a tangible interface using multimodal learning analytics. **Journal of Educational Data Mining**, International Working Group on Educational Data Mining, v. 7, n. 3, p. 89–116, 2015.

SCHNEIDER, Bertrand; PEA, Roy. Real-time mutual gaze perception enhances collaborative learning and collaboration quality. **International Journal of Computer-supported collaborative learning**, Springer, v. 8, n. 4, p. 375–397, 2013.

SHERIN, Bruce. A computational study of commonsense science: An exploration in the automated analysis of clinical interview data. **Journal of the Learning Sciences**, Taylor & Francis, v. 22, n. 4, p. 600–638, 2013.

SPIKOL, Daniel; CUKUROVA, Mutlu. *Multimodal Learning Analytics*, 2019.

SPIKOL, Daniel; RUFFALDI, Emanuele; DABISIAS, Giacomo; CUKUROVA, Mutlu. Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. **J. Comp. Assisted Learning**, v. 34, n. 4, p. 366–377, 2018. DOI: 10.1111/jcal.12263.

THOMPSON, Kate. Using micro-patterns of speech to predict the correctness of answers to mathematics problems: An exercise in multimodal learning analytics. *In: ACM. PROCEEDINGS of the 15th ACM on International conference on multimodal interaction*. [S.l.: s.n.], 2013. P. 591–598.

UPTON, Graham; COOK, Ian. **A dictionary of statistics 3e**. [S.l.]: Oxford university press, 2014.

WEINLAND, Daniel; RONFARD, Remi; BOYER, Edmond. Free viewpoint action recognition using motion history volumes. **Computer vision and image understanding**, Elsevier, v. 104, n. 2-3, p. 249–257, 2006.

WORSLEY, Marcelo. Multimodal Learning Analytics' Past, Present, and Potential Futures. *In: CROSSMMLA@ LAK*. [S.l.: s.n.], 2018.

WORSLEY, Marcelo. Multimodal Learning Analytics' Past, Present, and, Potential Futures, 2017.

WORSLEY, Marcelo; ABRAHAMSON, Dor; BLIKSTEIN, Paulo; GROVER, Shuchi; SCHNEIDER, Bertrand; TISSENBAUM, Mike. Situating multimodal learning analytics. *In: INTERNATIONAL SOCIETY OF THE LEARNING SCIENCES (ISLS). 12TH International Conference of the Learning Sciences: Transforming Learning, Empowering Learners, ICLS 2016. [S.l.: s.n.], 2016. P. 1346–1349.*

WORSLEY, Marcelo; BLIKSTEIN, Paulo. Analyzing engineering design through the lens of computation. **Journal of Learning Analytics**, ERIC, v. 1, n. 2, p. 151–186, 2014.

WORSLEY, Marcelo; BLIKSTEIN, Paulo. Towards the development of multimodal action based assessment. *In: ACM. PROCEEDINGS of the third international conference on learning analytics and knowledge. [S.l.: s.n.], 2013. P. 94–101.*

WORSLEY, Marcelo; BLIKSTEIN, Paulo. What's an Expert? Using Learning Analytics to Identify Emergent Markers of Expertise through Automated Speech, Sentiment and Sketch Analysis. *In: EDM. [S.l.: s.n.], 2011. P. 235–240.*

WÖRTWEIN, Torsten; CHOLLET, Mathieu; SCHAUERTE, Boris; MORENCY, Louis-Philippe; STIEFELHAGEN, Rainer; SCHERER, Stefan. Multimodal public speaking performance assessment. *In: PROCEEDINGS of the 2015 ACM on International Conference on Multimodal Interaction. [S.l.: s.n.], 2015. P. 43–50.*

YORK, Dustin. **Investigating a Relationship between Nonverbal Communication and Student Learning**. 2013. Tese (Doutorado) – Lindenwood University.

ZHANG, Zhengyou. Microsoft kinect sensor and its effect. **IEEE multimedia**, IEEE, v. 19, n. 2, p. 4–10, 2012.

ZHAO, Haiyong; LIU, Zhijing; ZHANG, Hao. Recognizing Human Activities Using Non-linear SVM Decision Tree. *In: CHEN, Ran (Ed.). **Intelligent Computing and Information Science**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. P. 82–92.*

ZHOU, Jianlong; HANG, Kevin; OVIATT, Sharon; YU, Kun; CHEN, Fang. Combining empirical and machine learning techniques to predict math expertise using pen signal

features. *In*: PROCEEDINGS of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge. [S.l.: s.n.], 2014. P. 29–36.

ZHU, Gaoxia; XING, Wanli; POPOV, Vitaliy. Uncovering the sequential patterns in transformative and non-transformative discourse during collaborative inquiry learning. **The Internet and Higher Education**, Elsevier, v. 41, p. 51–61, 2019.

APÊNDICE A – PUBLICAÇÕES

A.1 PUBLICAÇÕES

1. ROQUE, Felipe et al. Using Depth Cameras to Detect Patterns in Oral Presentations: A Case Study Comparing Two Generations of Computer Engineering Students. *Sensors*, v. 19, n. 16, p. 3493, 2019.
2. ROQUE, Felipe et al. Encontrando os padrões sequenciais em apresentações orais de estudantes utilizando Sequential Pattern Mining. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2019. p. 1896.
3. ROQUE, Felipe et al. Evolução das posturas corporais de estudantes em apresentações de seminários ao longo do semestre: uma análise utilizando dados multimodais e técnicas de clusterização. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2018. p. 1483.
4. ROQUE, Felipe Vieira et al. Using Multimodal Data to Find Patterns in Student Presentations. In: 2018 XIII Latin American Conference on Learning Technologies (LACLO). IEEE, 2018. p. 256-263.