

iscte

INSTITUTO
UNIVERSITÁRIO
DE LISBOA

A influência dos métodos de avaliação nos resultados dos alunos do ensino superior

Sónia Patrícia Gusmão Araújo

Mestrado em Informática e Gestão,

Orientadora:

Doutora Joana Martinho de Almeida Costa Pardal, Professora Auxiliar,

ISCTE-Instituto Universitário de Lisboa

Outubro, 2020



TECNOLOGIAS
E ARQUITETURA

A influência dos métodos de avaliação nos resultados dos alunos do ensino superior

Sónia Patrícia Gusmão Araújo

Mestrado em Informática e Gestão,

Orientadora:

Doutora Joana Martinho de Almeida Costa Pardal, Professora Auxiliar,

ISCTE-Instituto Universitário de Lisboa

Outubro, 2020

*Dedicada a todos os alunos que entram agora
para o ensino superior e que podem alcançar
melhores resultados*

[Esta página foi intencionalmente deixada em branco].

Agradecimento

A realização desta dissertação contou com o apoio e incentivo de diversas pessoas e sem as quais não era possível a sua realização e por isso estou grata.

Começando pela professora Joana Martinho Costa, minha orientadora, que me orientou nos momentos mais difíceis e desmotivantes, fazendo me sentir mais motivada e com vontade de continuar, apresentando sempre soluções e com a disponibilidade total, tornado assim o meu trabalho mais fácil.

Ao professor Pedro Ramos que disponibilizou a ferramenta desenvolvida pelo colega Hugo Miguel, a quem também agradeço, na qual consegui extrair parte dos dados necessários.

À professora Paula Odete Fernandes do IPB de Bragança que também me auxiliou bastante na parte estatística.

E por fim à minha família e amigos que estiveram sempre presentes quando precisei de ajuda. Em especial aos meus pais por financiarem o meu mestrado, ao meu namorado pela paciência e ajuda nos momentos de maior indecisão e aos meus amigos e colegas de curso, Pedro Pires e Pedro Carneiro por se terem disponibilizado para me ajudar em diversas questões.

A todos um sincero obrigada, sem vocês não seria possível.

[Esta página foi intencionalmente deixada em branco].

Resumo

O objetivo principal desta dissertação consistiu em analisar em que medida os métodos e instrumentos de avaliação têm influência nos resultados dos alunos do ensino superior bem como quais os métodos que têm maior impacto na aprovação, na qualidade e no sucesso escolar. Para atingir os objetivos propostos desenvolveu-se um estudo composto por duas fases. Na primeira fase, determinou-se o grau de influência de diversos métodos e instrumentos de avaliação nos alunos de Ensino Superior, utilizando a modelação de equações estruturais. Na segunda fase, foram aplicados os modelos de previsão RANDOMFOREST; KSVM; MLPE; MLP; MARS; RPART; MR; CV.GLMNET; RVM e XGBOOST, utilizando a metodologia KDD (Knowledge Discovery in Databases), a partir dos modelos gerados na primeira fase para prever os resultados dos alunos e as potenciais melhorias.

Para o desenvolvimento do estudo, foi utilizada uma amostra constituída por 797 médias das classificações de uma Instituição do Ensino Superior pertencente ao distrito de Lisboa em cursos das áreas de Informática e de Gestão entre 2013/2014 e 2017/2018.

Os resultados evidenciaram que (1) a utilização de laboratórios, projetos, apresentações, trabalhos individuais e de grupo são os elementos mais explicativos para o conjunto de todas as médias e para o conjunto que contém apenas as médias positivas; enquanto que (2) a utilização de uma avaliação em frequência combinada com o trabalho individual são os elementos mais explicativos para as médias elevadas.

Palavras-Chave: Métodos de avaliação, Ensino Superior, Educational Data Mining, Equações Estruturais, Sucesso Escolar

[Esta página foi intencionalmente deixada em branco].

Abstract

The main objective of this dissertation was to analyse the extent to which the assessment methods and instruments influence student's results in higher education as well as which methods have the greatest impact on approval and quality of academic success. To achieve the proposed objectives, a two-step study was developed. In the first stage, the degree of influence of several methods and instruments of assessment on higher education students was determined by using structural equations modelling. In the second phase, the prediction models RANDOMFOREST; KSVM; MLPE; MLP; MARS; RPART; MR; CV.GLMNET; RVM, and XGBOOST were applied using the KDD (Knowledge Discovery in Databases) methodology from the first phase's models to predict student outcomes and potential improvements.

In the development of the study, it was used a sample of 797 averages calculated from the grades of a Higher Education Institution located in Lisbon district. The courses of the sample belonged to the areas of Informatics and Management between the academic years 2013/2014 and 2017/2018.

The results showed that (1) the use of laboratories, projects, oral presentations, individual and group work are the most explanatory elements for the set of all averages and for the set of only positive averages; while (2) the use of the frequency evaluation method combined with individual work are the most explanatory elements for high averages.

Keywords: Academic Success, Assessment Methods, Higher Education, Educational Data Mining, Structural Equations

[Esta página foi intencionalmente deixada em branco].

Índice

Agradecimento.....	i
Resumo	iii
Abstract	v
Índice Figuras	ix
Índice Tabelas.....	xi
Lista de acrónimos	xiii
1. Introdução.....	1
1.1 Motivação	1
1.2 Objetivos	1
1.3 Metodologias	2
1.4 Questões de investigação e Hipóteses.....	2
1.5 Estrutura e Organização.....	3
2. Estado de Arte	5
2.1 Definição de avaliação e de aprendizagem	5
2.2 Tipos de avaliação	5
2.3 Métodos de avaliação	7
2.4 Data Mining aplicado ao ensino superior.....	9
3. Metodologia.....	13
3.1 Opções metodológicas	13
3.2 Descrição da recolha e tratamento dos dados	15
4. Resultados	19
4.1 Análise de dados	19
4.2 Modelação de equações estruturais	23
4.2.1 Estudo 1	24
4.2.2 Estudo 2	32
4.2.3 Estudo 3	35
4.2.4 Discussão	39
4.3 Educational Data Mining.....	39
4.3.1 Estudo 1	41
4.3.2 Estudo 2	44
4.3.3 Estudo 3	47
4.3.4 Estudo 4	50
4.3.5 Discussão	53
4.4 Estudo Complementar	54
5. Conclusão	59
Referências	61

ANEXOS	65
Anexo A – Tabela Coeficientes Stepwise Estudo 1 (Modelação de EE)	65
Anexo B – Tabela Coeficientes Stepwise Estudo 2 (Modelação de EE).....	66
Anexo C – Figuras dos gráficos de importância do Estudo 1 de EDM.....	67
Anexo D – Figuras dos gráficos de importância do Estudo 2 de EDM.....	68
Anexo E – Figuras dos gráficos de importância do Estudo 3 de EDM.....	69
Anexo F – Figuras dos gráficos de importância do Estudo 4 de EDM.....	70

Índice Figuras

Figura 1.1: Fluxograma	4
Figura 2.1.Métodos de avaliação mais utilizados, segundo o estudo de Pereira e Flores (2012).....	7
Figura 2.2.Métodos de avaliação menos utilizados, segundo o estudo de Pereira e Flores (2012).....	8
Figura 2.3 Principais áreas relacionadas com EDM (Romero e Ventura (2013)).....	12
Figura 3.1: Framework da aplicação de Miguel et al. (2020).....	16
Figura 3.2:Layout de pesquisa da Aplicação de Miguel et al. (2020)	16
Figura 3.3: Resultado da pesquisa.....	17
Figura 3.4:Organização dos dados (Excel).....	18
Figura 4.1: Modelo 1 Estudo 1	25
Figura 4.2: Modelo 2 Estudo1	27
Figura 4.3: Modelo 3 Estudo1	28
Figura 4.4: Modelo 4 Estudo1	29
Figura 4.5: Modelo 5 Estudo1	30
Figura 4.6: Modelo 6 Estudo1	31
Figura 4.7: Modelo 7 Estudo1	32
Figura 4.8: Modelo 1 Estudo 2	33
Figura 4.9: Modelo 2 Estudo2.....	34
Figura 4.10: Modelo 3 Estudo 2.....	35
Figura 4.11: Modelo 1 Estudo 3.....	37
Figura 4.12: Modelo 2 Estudo 3.....	38
Figura 4.13: Gráfico de Importância Estudo 1 randomForest.....	42
Figura 4.14:Gráfico de Importância Estudo 1 KSVM.....	42
Figura 4.15: Gráfico de Médias Modelos Estudo 1.....	43
Figura 4.16: Máximos e Mínimos Estudo 1	44
Figura 4.17: Gráfico de Importância Estudo 2 KSVM.....	45
Figura 4.18: Gráfico de Importância Estudo 2 MLPE	46
Figura 4.19: Gráfico de Médias Modelos Estudo 2.....	46
Figura 4.20: Máximos e Mínimos Estudo 2	47
Figura 4.21: Gráfico de Importância Estudo 3 KSVM.....	48
Figura 4.22:Gráfico de Importância Estudo 3 Rpart	49
Figura 4.23: Gráfico de Médias Modelos Estudo 3.....	49
Figura 4.24: Máximos e Mínimos Estudo 3	50
Figura 4.25: Gráfico de Importância Estudo 4 KSVM.....	51
Figura 4.26: Gráfico de Importância Estudo 4 randomForest.....	51
Figura 4.27: Gráfico de Médias Modelos Estudo 4.....	52
Figura 4.28: Máximos e Mínimos Estudo 4	52
Figura 4.29: Máximos e Mínimos Estudo 1 vs. Estudo 2	54
Figura 4.30: Gráfico Médias por grupo.....	56
Figura 4.31:Boxplot areas científicas	57
Figura 4.32: Gráfico Médias por ano letivo	58

[Esta página foi intencionalmente deixada em branco].

Índice Tabelas

Tabela 2.1: Avaliação Sumativa vs. Avaliação Formativa.....	6
Tabela 2.2: Técnicas de avaliação de sala de aula.....	9
Tabela 3.1: Fases e Etapas do processo de KDD (Pimentel & Omar, 2006)	14
Tabela 3.2: Frequências.....	18
Tabela 4.1: Teste de Kolmogorov-Smirnov de Uma Amostra.....	20
Tabela 4.2: Assimetria e Curtose	21
Tabela 4.3: Correlação	22
Tabela 4.4: Frequências PesoE.....	23
Tabela 4.5: Pesos Regressão Modelo 1 Estudo 1	24
Tabela 4.6: Coeficientes e Estatísticas de Colinearidade	25
Tabela 4.7: Sumário dos Modelos	26
Tabela 4.8: Coeficientes Stepwise Estudo 1.....	27
Tabela 4.9: Pesos Regressão Modelo 2 Estudo 1	28
Tabela 4.10: Pesos Regressão Modelo 3 Estudo 1	28
Tabela 4.11: Pesos Regressão Modelo 4 Estudo 1	29
Tabela 4.12: Pesos Regressão Modelo 5 Estudo 1	30
Tabela 4.13: Pesos Regressão Modelo 6 Estudo 1	30
Tabela 4.14: Pesos Regressão Modelo 7 Estudo 1	31
Tabela 4.15: Pesos Regressão Modelo 8 Estudo 1	32
Tabela 4.16: Frequências (Valores média ≥ 10)	32
Tabela 4.17: Pesos Regressão Modelo 1 Estudo 2.....	33
Tabela 4.18: Coeficientes Stepwise Estudo 2.....	34
Tabela 4.19: Pesos Regressão Modelo 2 Estudo2.....	34
Tabela 4.20: Pesos Regressão Modelo 3 Estudo2.....	35
Tabela 4.21: Coeficientes Estudo 3	36
Tabela 4.22: Variáveis Excluídas	36
Tabela 4.23: Pesos Regressão Modelo 1 Estudo 3	37
Tabela 4.24: Coeficientes Stepwise Estudo 3.....	38
Tabela 4.25: Pesos Regressão Modelo 2 Estudo 3	38
Tabela 4.26: Estudo 1 Métricas	41
Tabela 4.27: Estudo 2 Métricas	45
Tabela 4.28: Estudo 3 Métricas	48
Tabela 4.29: Estudo 4 Métricas	51
Tabela 4.30: Resumo do processamento de casos	55
Tabela 4.31: Teste à Normalidade.....	55
Tabela 4.32: Assimetria e Curtose	56
Tabela 4.33: Teste Estatístico Kruskal-Wallis	57

[Esta página foi intencionalmente deixada em branco].

Lista de acrónimos

EDM	Educational Data Mining
EE	Equações Estruturais
FUC	Ficha de Unidade Curricular
KDD	Knowledge Discovery in Databases
KMA	Knowledge Monitoring Accuracy
KMB	Knowledge Monitoring Bias
MEE	Modelação de Equações Estruturais
UC	Unidade curricular

[Esta página foi intencionalmente deixada em branco].

1. Introdução

A avaliação das aprendizagens é um tema muito debatido entre a comunidade académica pela sua pertinência no sucesso dos estudantes. Neste sentido importa abordar com maior profundidade de que forma os métodos de avaliação podem afetar as médias dos alunos, uma vez que este é um ponto fulcral no rendimento académico.

Fernandes, Rodrigues e Nunes (2012) afirmam que, no que trata à investigação da avaliação no ensino superior, os estudos são insuficientes e que por isso, existe a necessidade de compreender questões cruciais que ajudassem a melhorar a avaliação.

O presente capítulo pretende fazer uma introdução à motivação, aos objetivos, às questões de investigação, às hipóteses, à estrutura, à organização desta dissertação e às metodologias utilizadas.

1.1 Motivação

A escolha deste tema começou com uma simples questão já antiga que me intrigava e à qual desconhecia a resposta: será que existe alguma relação entre os métodos de avaliação e os resultados dos alunos? Questão essa que deu origem a muitas outras questões como por exemplo; será que numa avaliação por projetos os alunos têm melhores notas e consequentemente aprendem mais e melhor? Ou será que a avaliação por exame é o mais indicado para avaliar conhecimentos?

Segundo Maffezzolli e Moraes Pereira (2020) as escolhas dos métodos de avaliação utilizados pelos docentes podem aumentar ou diminuir as oportunidades para que os alunos demonstrem o que aprenderam.

Para tentar resolver este problema e alinhando com gosto especial pela área de Data Mining surgiu a ideia de tentar perceber através de técnicas de Data Mining a influência que os métodos de avaliação podem ter nos resultados dos alunos universitários.

1.2 Objetivos

O objetivo principal do estudo é perceber em que medida os métodos de avaliação têm influência nos resultados dos alunos do ensino superior. E por consequência perceber quais os métodos com maior impacto na aprovação e sucesso escolar através da utilização de modelação

de equações estruturais. Após este objetivo ser alcançado será apresentado um estudo na área “Educational Data Mining” recorrendo à metodologia KDD com o método de partição k-fold e estudando dez modelos de previsão para prever o comportamento das médias dos alunos, caso os métodos de avaliação utilizados fossem os extraídos dos modelos das equações estruturais.

Esta dissertação está inserida no contexto da educação no ensino superior e de que modo a modelação de equações estruturais e as tecnologias de Data Mining podem ajudar a melhorar o ensino, mais concretamente a gestão dos métodos de avaliação nas diversas unidades curriculares.

1.3 Metodologias

A metodologia utilizada para a recolha dos dados foi feita através da utilização de uma ferramenta desenvolvida por Miguel (2019) e pelas Fichas das Unidades Curricular onde foi possível recolher o método de avaliação de cada Unidade Curricular.

Para a modelação de equações estruturais foi utilizado o programa SPSS V26 e o AMOS V26 para a modelação e para a área de Educational Data Mining (EDM) foi utilizada a ferramenta RStudio bem como dez modelos de previsão aplicados a problemas de regressão.

1.4 Questões de investigação e Hipóteses

A questão de investigação principal é a seguinte: “Será que existe relação entre as variáveis “método de avaliação” e os resultados dos alunos?”

Questão essa que se divide em duas hipóteses estatísticas:

H_0 : Não existe relação entre os métodos de avaliação e os resultados dos alunos.

H_a : Existe relação entre os métodos de avaliação e os resultados dos alunos.

Após o estudo da primeira questão de investigação será analisado de que forma os métodos de avaliação podem afetar a variável média, isto é, quanto é que estas variáveis são capazes de explicar a variável média recorrendo à modelação de equações estruturais. E por fim prever como seriam as médias nas diferentes Unidades Curriculares caso fossem utilizados os métodos e instrumentos de avaliação mais adequados a cada finalidade, nomeadamente, o mais explicativo do total de médias registadas, o mais explicativo das médias positivas e o mais explicativo das médias superiores a 14 valores.

Extraiu-se a partir daqui uma segunda questão de investigação: “Em que medida as variáveis “método de avaliação” podem afetar a variável média?” E uma terceira questão: “Como é que as médias podem ser afetadas se utilizarmos apenas alguns tipos de avaliação?”

Uma vez que os dados podem ser mais trabalhados e estudados, será ainda apresentado um subcapítulo sobre as áreas das Unidades Curriculares (Código da Cadeira) e as médias das Unidades Curriculares sendo as hipóteses estatísticas as seguintes:

H_0 : Não existem diferenças entre os grupos “código da cadeira” relativamente à média.

H_a : Existem diferenças entre os grupos “código da cadeira” relativamente à média.

1.5 Estrutura e Organização

Esta dissertação está organizada da seguinte forma: inicialmente apresenta-se um capítulo denominado por Estado de Arte que contém a informação literária necessária para contextualizar o trabalho. De seguida surge o capítulo Metodologias onde é descrito o processo de recolha dos dados e as metodologias utilizadas para a modelação de equações estruturais bem como as metodologias e processos utilizados no campo do EDM.

No quarto capítulo apresentam-se os resultados obtidos pelo estudo. Este capítulo está dividido em três subcapítulos sendo o primeiro referente à modelação de equações estruturais, o segundo o estudo referente à área de EDM e um terceiro subcapítulo que contém um estudo sobre em que áreas os alunos têm melhores resultados e se existe ou não diferenças entre as áreas das Unidades Curriculares e os resultados. Importa referir que os dois primeiros subcapítulos contêm uma secção denominada por “Discussão” com pequenas conclusões sobre os temas abordados (Modelação de Equações Estruturais (MEE) e Educational Data Mining (EDM)). O penúltimo capítulo apresentado corresponde à Conclusão e o último lista as referências bibliográficas.

Para facilitar o seguimento do estudo e a sua organização, apresenta-se o seguinte fluxograma que descreve as fases e as principais tomadas de decisão que serão descritas ao longo desta dissertação (Figura 1.1).

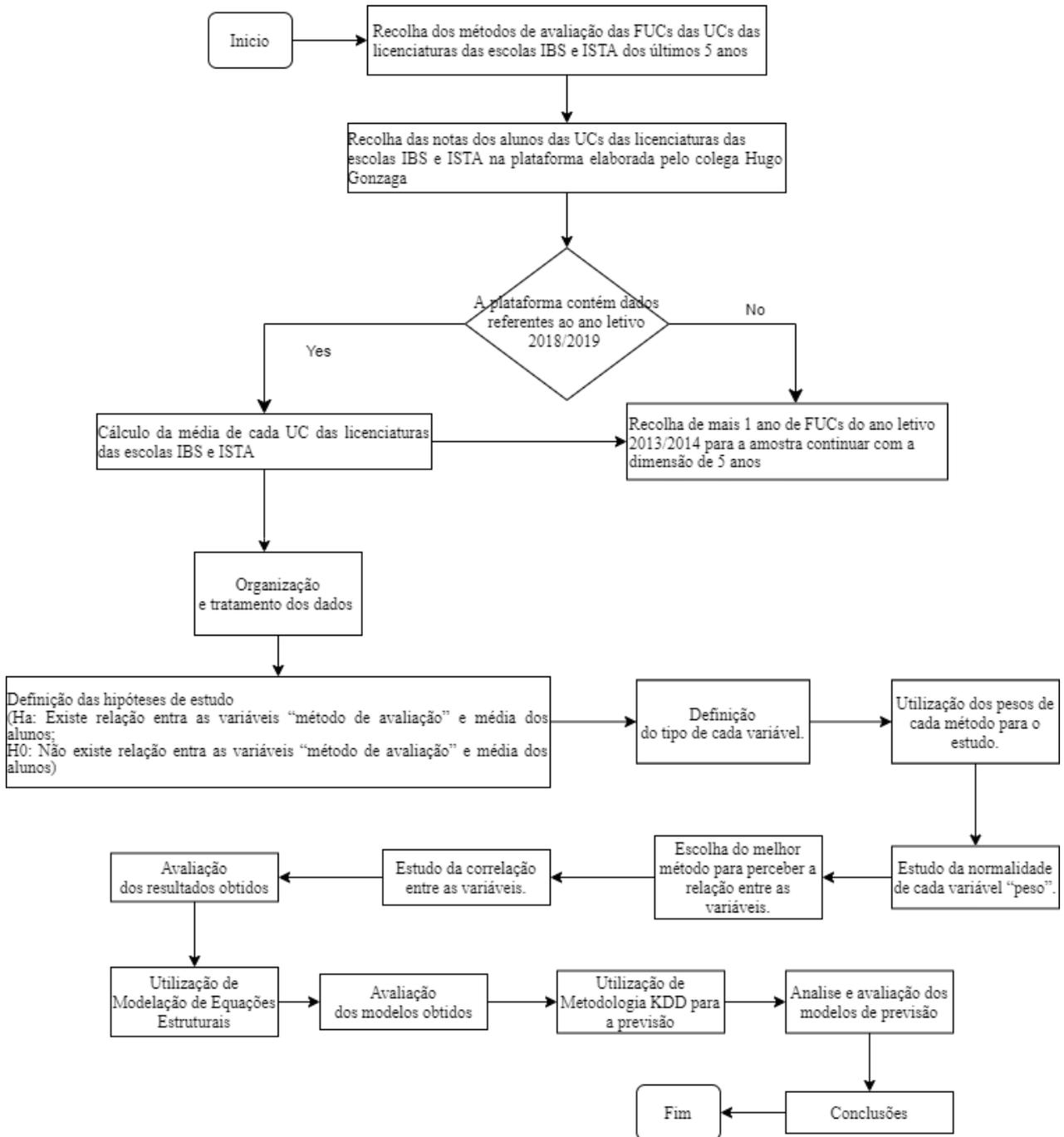


Figura 1.1: Fluxograma

2. Estado de Arte

2.1 Definição de avaliação e de aprendizagem

A avaliação é um processo de verificação de objetivos, em que, o que os alunos produzem a nível escolar é comparado a um modelo. Este processo contribui para a eficácia do ensino, uma vez que consiste na observação e interpretação dos seus resultados. É no processo de ensino-aprendizagem que surge a avaliação, funcionando como um mecanismo que verifica se os objetivos pretendidos foram efetivamente atingidos (Rosado & Silva, 2010). Trotter (2006) refere que, a avaliação é um termo genérico utilizado para um conjunto de processos que medem o resultado da aprendizagem dos alunos. E que uma das melhores formas de progredir na aprendizagem dos alunos é alterar o método de avaliação que está implementado.

Muitos autores referem que a avaliação e a aprendizagem estão interligadas. É o caso de Pereira e Flores (2012), que afirma que, de entre as treze variáveis estudadas, nomeadamente, a aprendizagem, os testes, a verificação de conhecimentos, o sucesso, as notas/classificação, a reflexão, a participação, a ansiedade, a certificação, a ajuda, o conflito, o receio/medo, a imposição e a injustiça, a variável que mais se destaca para os alunos é a variável “aprendizagem”. Segundo Barbosa (2008) a avaliação é uma tarefa necessária do trabalho docente, que deve acompanhar passo a passo o processo de ensino e aprendizagem. A mesma autora refere ainda que a “avaliação é uma reflexão sobre o nível de qualidade do trabalho escolar tanto do professor como dos alunos” (Barbosa, 2008, p.1).

Em 2012, Fernandes, Rodrigues e Nunes (2012) afirmavam que, no que trata à investigação da avaliação e das aprendizagens no ensino superior, os estudos eram ainda insuficientes e que por isso, existia uma necessidade de compreender questões cruciais que ajudassem a melhorar a avaliação e a aprendizagem neste nível de ensino.

2.2 Tipos de avaliação

A avaliação divide-se em duas principais vertentes: a avaliação formativa e a avaliação sumativa (Fernandes & Fialho, 2012; Fernandes et al., 2012; Miranda, 1982). No entanto, alguns autores que consideram também a avaliação do tipo prognóstico e diagnóstico (e.g. Barbosa, 2008; Rosado & Silva, 2010; Leitão, 2013).

A avaliação formativa tem presente a relação entre a pedagogia e o aluno. Ou seja, acompanha o percurso e adequa constantemente os métodos que o aluno precisa para progredir e

melhorar. Por este motivo, este modelo de avaliação é considerado o próprio ensino, uma vez que ao ensinar, o professor está a aplicar os seus métodos educativos que permitem ao aluno ter a plena consciência do seu ponto de aprendizagem e como se deve ajustar para atingir os seus objetivos. Esta avaliação foca-se principalmente nas atividades e tem como finalidade facilitar a aprendizagem (Leitão, 2013). Em suma, é importante que os alunos respondam ao que é perguntado sem a atribuição de uma classificação como consequência da sua resposta, de forma a diminuir os níveis elevados de stress e ansiedade (Rosado & Silva 2010), habitualmente detetados em situações de avaliação sumativa.

Fernandes, Rodrigues e Nunes (2012) fazem referência a um artigo de revisão de literatura de Black e Wiliam (1998) , sobre práticas de avaliação formativa onde evidenciam que: a) a prática sistemática de avaliação formativa melhora a aprendizagem dos alunos; b) os alunos que mais beneficiam da avaliação formativa são aqueles que apresentam maiores dificuldades e c) os alunos que frequentam aulas em que a avaliação formativa predomina, obtêm melhores resultados nos exames de avaliação externa do que aqueles que frequentam aulas em que o modelo de avaliação predominante é de natureza sumativa.

A avaliação sumativa tem como função verificar e quantificar aquilo que os alunos retiveram. É por isso uma avaliação terminal e retrospectiva. É aplicada no final do processo de aprendizagem e, tal como a avaliação diagnóstica, centra-se no que os alunos são capazes de produzir. Para além disso, situa os alunos em diferentes patamares de acordo com o resultado obtido, sendo por isso, classificativa (Leitão, 2013).

Na Tabela 2.1 observam-se os objetivos e diferenças entre avaliações sumativa e formativa, de acordo com Chism, Angelo e Cross (1995):

Tabela 2.1: Avaliação Sumativa vs. Avaliação Formativa

Sumativa	Formativa
<u>Objetivo</u> : julgar o desempenho individual do aluno e atribuir notas	<u>Objetivo</u> : informar os professores e melhorar a aprendizagem; utilizando feedback
<u>Exemplos</u> : quizzes, testes, exames; trabalhos de pesquisa, relatórios de laboratório, trabalhos de casa.	<u>Exemplos</u> : fazer perguntas, ouvir as perguntas e comentários dos alunos, perceber a linguagem corporal e expressões faciais.

Fernandes e Fialho (2012) referem que vários autores consideram a não existência do modelo de avaliação formativa no ensino superior (e.g. Gijbels, Segers, & Struyf, 2008; Garcia, 2009. Reconhecem ainda que consideram existir a necessidade da utilização deste tipo de prática

avaliativa referindo que a avaliação baseia-se quase exclusivamente na transmissão de conhecimentos que são avaliados através de testes finais ((Bryan & Clegg, 2006).

A avaliação diagnóstica tem como principal objetivo chegar a um diagnóstico, isto é, enumerar os saberes do aluno, determinar o ponto de situação do nível do seu conhecimento e orientá-lo. Um diagnóstico é sempre associado a um início de procedimento e por isso deve ser sempre aplicado antes de qualquer processo educativo. Este tipo de avaliação serve para identificar os pontos fortes e fracos de cada aluno, permitindo assim ao docente selecionar a metodologia que mais se adequa ao aluno (Leitão, 2013).

2.3 Métodos de avaliação

Segundo Pereira e Flores (2012) os métodos de avaliação mais utilizados pelos docentes, de acordo com a opinião dos alunos, são: as apresentações orais na aula em grupo, os testes ou exames escritos, os relatórios realizados em grupo, a resolução de trabalhos práticos em grupo, o trabalho de projeto realizado em equipa e ainda, os relatórios individuais, tal como se pode observar na Figura 2.1.

Métodos de Avaliação	N	Nada Utilizados	Pouco Utilizados	Bastante Utilizados	Muito Utilizados	Total
Apresentações orais na aula em grupo	254	0,8%	8,3%	48,4%	42,5%	100%
Testes ou exames escritos	253	0,4%	11,1%	17,8%	70,8%	100%
Relatórios realizados em grupo	251	1,2%	12%	55,8%	31,1%	100%
Resolução de trabalhos práticos/ Trabalhos Experimentais em grupo	252	0,8%	13,1%	50,4%	35,7%	100%
Trabalho de projeto realizado em equipa	254	2,8%	23,6%	47,2%	26,4%	100%
Relatórios individuais	253	7,5%	43,1%	36,8%	12,6%	100%

Figura 2.1. Métodos de avaliação mais utilizados, segundo o estudo de Pereira e Flores (2012)

Os resultados mostram que a maior parte dos métodos de avaliação utilizados remetem para o trabalho em grupo. No entanto, os testes ou exames destacam-se como um método muito utilizado pelos professores (Pereira & Flores, 2012). Esta opção poderá estar relacionada com potencialidades particulares destes métodos de avaliação tais como a hierarquização, a promoção da equidade, a garantia da qualidade do ensino, a credibilidade do sistema, o cumprimento dos programas, a herança cultural e os efeitos de adaptação, na escola, no professor e no aluno (Leal, 1997).

Em relação aos métodos de avaliação menos utilizados, como se constata na Figura 2.2, destacaram-se: os ensaios individuais, os ensaios em grupo, os testes ou exames orais, as recensões críticas de textos individuais, os portefólios individuais, as recensões críticas de textos em grupo, os portefólios coletivos, as apresentações orais individuais, os mini-testes, as reflexões escritas individuais, o trabalho de projeto individual e, por último, a resolução de trabalhos experimentais individuais (Pereira & Flores; 2012).

Métodos de avaliação	N	Nada Utilizados	Pouco Utilizados	Bastante Utilizados	Muito Utilizados	Total
Ensaio individuais	254	56,7%	37%	4,7%	1,6%	100%
Ensaio realizados em grupo	254	48,8%	33,1%	14,2%	3,9%	100%
Testes ou exames orais	254	21,3%	59,8%	14,2%	4,7%	100%
Recensões críticas de textos individuais	252	43,7%	34,5%	12,3%	9,5%	100%
Portefólios individuais	254	35,4%	41,7%	18,5%	4,3%	100%
Recensões críticas de textos em grupo	252	44%	32,9%	18,3%	4,8%	100%
Portefólios coletivos	252	41,7%	29%	21,8%	7,5%	100%
Apresentações orais na aula individuais	253	16,2%	53,4%	21,7%	8,7%	100%
Mini-testes individuais	253	36,8%	30,8%	14,6%	17,8%	100%
Reflexões escritas individuais	254	24,4%	34,3%	22%	19,3%	100%
Trabalho de projeto realizado individualmente	254	11,4%	46,5%	34,6%	7,5%	100%
Resolução de trabalhos práticos/ Trabalhos Experimentais individuais	253	6,7%	44,3%	32,4%	16,6%	100%

Figura 2.2. Métodos de avaliação menos utilizados, segundo o estudo de Pereira e Flores (2012)

Para obter esta informação, as autoras realizaram um estudo com 254 inquiridos, dos quais 165 alunos de mestrado integrado e 89 alunos de licenciatura. Importa referir que todos os alunos eram do 3º ano da Universidade do Minho (Pereira & Flores (2012). No mesmo estudo, as autoras afirmam que a grande maioria dos alunos considera que a avaliação seria mais justa se fossem utilizados pelo menos dois métodos de avaliação diferentes.

Pereira e Flores (2012) fazem referência a Rasmden (2003) onde o autor alerta que é difícil combinar resultados de aprendizagem a partir de diferentes métodos de avaliação, dando como exemplo que a classificação de um trabalho prático ou um projeto será diferente da classificação atingida pelo mesmo aluno num teste ou exame. Por esse motivo é importante que os sistemas de avaliação se tornem criativos no que respeita as práticas de avaliação adotadas, devendo ser possível utilizar métodos diversificados e inovadores.

Gaulden (2010, citado por (Chism et al., 1995), resume as diversas técnicas de avaliação em sala de aula e agrupa-as por categorias e subcategorias. Sendo estas (ver Tabela 2.2):

Tabela 2.2: Técnicas de avaliação de sala de aula

Conhecimentos e competências relacionadas com o curso	<p>Avaliar o conhecimento anterior, o recordar e a compressão;</p> <p>Avaliar a capacidade de análise e o pensamento crítico;</p> <p>Avaliar a capacidade de sintetizar e o pensamento criativo;</p> <p>Avaliar a capacidade de resolução de problemas;</p> <p>Avaliar a capacidade na aplicação e no desempenho.</p>
Atitudes, valores e autoconsciência dos alunos	<p>Avaliar a sensibilidade dos alunos para as atitudes e valores dos mesmos;</p> <p>Avaliar a autoconsciência dos alunos;</p> <p>Avaliar a aprendizagem relacionada com o curso e habilidades de estudo, estratégias e comportamentos;</p>
Reações dos alunos à instrução	<p>Avaliar as reações dos alunos aos professores e ao ensino;</p> <p>Avaliar as reações dos alunos às atividades, trabalhos e materiais da aula;</p>

Através dos subcapítulos 2.1 e 2.2 é possível observar que existem diferentes tipos de avaliação, como a sumativa e formativa e diferentes métodos de avaliação como por exemplo, mini testes, trabalhos de grupo, testes, exames, entre muitos outros. No entanto, como um dos objetivos do estudo é analisar quais as combinações mais explicativas, será necessário utilizar mecanismos de Data Mining, mais concretamente relativos à área EDM (*Educational Data Mining*). O subcapítulo 2.4 descreve esta técnica, que será utilizada a partir do Capítulo 3.

2.4 Data Mining aplicado ao ensino superior

Segundo Baepler e Murdoch (2010), a análise académica e o Data Mining educativo estão a produzir rapidamente novas possibilidades de recolha, análise e apresentação de dados dos alunos. Os algoritmos tradicionais de Data Mining não podem ser aplicados diretamente em problemas educativos, uma vez que esses problemas podem ter um objetivo e uma função específica (Dutt,

Ismail, & Herawan, 2017). Por esse motivo é necessário que um algoritmo de pré-processamento de dados tenha que ser aplicado primeiro e só depois aplicar os métodos de Data Mining.

Foi em 2004 que surgiu o termo “*data mining educativo*” (EMD), no entanto, o Data Mining aplicado ao ensino surgiu em 1995. Para Baepler e Murdoch (2010), o Data Mining é semelhante a um grande campo de dados a partir do qual se podem desenterrar todos os tipos de informação insensíveis que, quando manipuladas por estas técnicas, podem apresentar conhecimentos úteis.

Data Mining aplicado ao ensino superior, pode significar perceber um comportamento, em particular num Sistema de Gestão de Cursos. Por exemplo, se um aluno que coloca um “x” de comentários num fórum on-line pode levar a um ganho de “y” na nota final desse aluno (Baepler & Murdoch, 2010).

Para os autores Castro, Vellido, Nebot, e Mugica (2007), a contante crescente da Internet permitiu que o ensino à distância online se tornasse muito mais utilizado. Este tipo de sistemas gera uma quantidade exponencial de dados, e grande parte desta informação tem o potencial de se tornar um novo conhecimento para melhorar todas as instâncias de e-learning. Os métodos de Data Mining devem permitir a extração desse conhecimento.

Ao longo dos anos, têm sido publicados diversos artigos sobre a investigação académica nesta área. Do ponto de vista dos problemas de Data Mining, os autores observaram que estes são dominados pela pesquisa sobre classificação e agrupamento. Este resultado não é surpreendente para os autores, dada a variedade e ampla disponibilidade de métodos, técnicas e ferramentas de software de Data Mining para ambos (Castro, Vellido, Nebot, & Mugica 2007).

Romero, Ventura e Garcia (2008) sustentam que Data Mining é uma área multidisciplinar na qual existem vários paradigmas de computação, tais como: a construção de árvores de decisão, a indução de regras, redes neurais artificiais, a aprendizagem baseada em instâncias, a aprendizagem bayesiana, a programação lógica e os algoritmos estatísticos. Referem ainda alguns métodos mais utilizados em Data Mining sendo estes: a estatística, a visualização, o agrupamento, a classificação, a associação de regras de mineração, a mineração de padrões sequenciais e o TextMining,

O conhecimento obtido pela utilização de técnicas de Data Mining, pode ser orientado aos alunos, de modo a recomendar atividades que favoreçam a sua aprendizagem, ou orientado aos professores, de forma a obter feedback mais objetivo para a educação, a avaliar a estrutura do conteúdo do curso e eficácia no processo de aprendizagem, a classificar os alunos em grupos com base nas suas necessidades de orientação, a encontrar padrões regulares, bem como padrões irregulares do aluno e a encontrar os erros cometidos com mais frequência por parte dos alunos.

Baker, Isotani e Carvalho (2011) afirmam que é possível aplicar modelos de Data Mining para verificar a relação entre a abordagem pedagógica e a aprendizagem dos alunos. Através desta informação o professor conseguiria compreender se a sua abordagem está realmente a ajudar os alunos e poderia desenvolver novos métodos de ensino mais eficazes.

A EDM é definida como a área de pesquisa que tem como principal objetivo o desenvolvimento de métodos para explorar conjuntos de dados recolhidos em diversos ambientes educativos. Para que seja possível compreender de maneira mais eficaz e adequada os alunos (e.g.: como os alunos aprendem, os fatores que influenciam a sua aprendizagem). É possível identificar em que situação um tipo de abordagem instrutiva proporciona melhores benefícios educativos ao aluno. É também possível verificar se o aluno está desmotivado ou confuso e, desse modo, personalizar os métodos de ensino para oferecer melhores condições de aprendizagem (Baker, Isotani & Carvalho, 2011).

Baker, Isotanie e Carvalho (2011) salientam que existe uma falta de independência estatística nos tipos de dados encontrados ao recolher informações no contexto educativo. Por esse motivo, diversos algoritmos e ferramentas utilizadas na área de Data Mining não podem ser aplicadas para analisar dados educativos sem modificação, não oferecendo apoio para validação cruzada entre os dados no nível do aluno.

A validação cruzada permite verificar a correção de um modelo criado a partir da análise de dados de treino (*training data*). Essa validação oferece uma estimativa de como o modelo se irá comportar ao analisar um novo conjunto de dados. A validação cruzada ao nível de aluno é fundamental em dados educativos, pois existe uma grande quantidade de dados por aluno e as conclusões obtidas ao utilizar métodos de Data Mining precisam de garantir que o modelo encontrado possa ser utilizado para inferir o comportamento ou a aprendizagem dos alunos (Baker, Isotani & Carvalho, 2011).

Os autores Baker, Isotanie e Carvalho (2011) apresentam uma introdução dos tópicos mais interessantes na área de EDM, como, i) a predição, que inclui a classificação, a regressão e a estimativa de densidade; ii) o clustering; iii) a mineração de relações, onde se insere a mineração de regras de associação, iv) a mineração de correlações e a mineração de padrões sequenciais; v) a destilação de dados para facilitar decisões humanas; e por fim, vi) descobertas com modelos.

Na predição, a área aplicada nesta dissertação, o objetivo consiste em desenvolver modelos que deduzam aspetos específicos dos dados, conhecidos como *predicted variables*, através da análise e fusão dos diversos aspetos encontrados nos dados, chamadas *predictor variables*. A estimativa de densidade é raramente utilizada na EDM devido a falta de independência estatística dos dados (Baker, Isotani, & Carvalho, 2011).

Baker, Isotani e Carvalho (2011) afirmam que existe atualmente uma pesquisa crescente na área de EDM, em particular nos Estados Unidos, Canadá e Espanha. Os autores destacam uma situação de pesquisa nos Estados Unidos onde os dados das escolas estão a ser disponibilizados aos investigadores através de bases de dados públicas como, por exemplo *National Center for Education Statistics*¹. Estes recursos permitem que os investigadores possam, mais facilmente, obter grandes quantidades de dados reais e relevantes para realizarem análises utilizando técnicas provenientes de EDM. Os investigadores que utilizam estes dados podem conduzir investigações com alta validade ecológica, ou seja, os resultados podem ser utilizados no contexto escolar enquanto se evita custos tradicionais da pesquisa e recolha de dados nesta área.

Dutt, Ismail e Herawan (2017) indicam diferentes estudos e investigadores na área de EDM, como por exemplo Bresfelean, Ghisoiu e Comes (2008) e de Vandamme, Meskens, e Superby (2007) que foram capazes de agrupar alunos com base no desempenho académico nos exames; Romero, Ventura e Garcia (2008) que elaboraram diversos estudos sobre a aplicação de Data Mining nas instituições educativas tradicionais para sistemas de gestão de aprendizagem baseados na web; Peña-Ayala (2014) concluiu na sua investigação que maioria dos trabalhos de investigação em EDM focada em três tipos de sistemas educativos, sendo estes, as tarefas educativas, os métodos e os algoritmos; Zaïane (2001) investigaram a aplicação de regras de associação e de clustering para apoiar o desenvolvimento e melhoramento dos sistemas de e-learning.

Romero e Ventura (2013) apresentam um esquema das principais áreas relacionadas com o EDM (ver Figura 2.3):

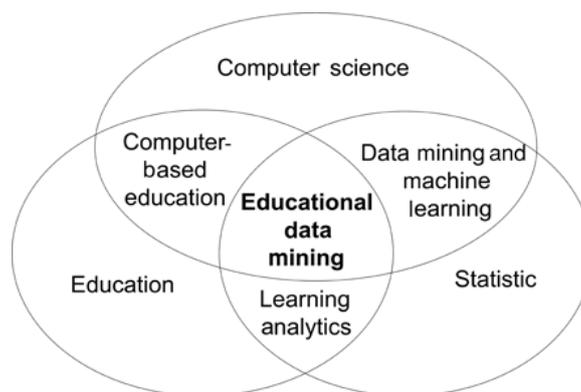


Figura 2.3 Principais áreas relacionadas com EDM (Romero e Ventura (2013))

¹<https://nces.ed.gov/>

Gil, Martins, Moro, e Costa (2020) apresentaram um estudo que teve como principal objetivo, aplicar técnicas de Data Mining a um conjunto de dados académicos fornecidos por uma Instituição de Ensino Superior, para extrair informação significativa e assim ajudarem a aumentar a taxa de sucesso académico. No estudo, os autores verificaram que a nota de entrada, o tempo de interrupção dos estudos e a idade dos alunos são outros fatores preditivos do sucesso académico no primeiro ano do ensino superior.

Estes trabalhos foram, escolhidos porque utilizam a nomenclatura mais aproximada da amostra que se irá trabalhar. Ir-se-á optar pela utilização das notas dos alunos universitários em avaliação contínua. Foi possível através da revisão de literatura, perceber-se que existem diversos artigos na área de EDM e que é uma área cada vez mais estudada. No entanto, não foram encontrados artigos que estudem exatamente o que é pretendido investigar nesta dissertação. A questão de investigação principal é tentar compreender de que forma os diferentes métodos de avaliação podem influenciar as médias dos alunos no ensino superior. Para isso, recorrer-se-á à utilização da modelação de equações estruturais e posteriormente a utilização de modelos de previsão para compreender como podem evoluir as médias dos alunos.

3. Metodologia

3.1 Opções metodológicas

Para responder às questões de investigação, foram consideradas três vertentes de análise. Na primeira, o objetivo principal foi compreender em que medida os valores de média compreendidos entre 0 e 20 podem ser explicados. Na segunda, o objetivo fundamental foi perceber que variáveis explicavam notas iguais ou superiores a 10, isto é, médias onde os alunos teriam sucesso. E por fim, a terceira vertente consistiu em perceber quais as variáveis que conseguiam explicar bons resultados, sendo que os bons resultados foram considerados como médias iguais ou superiores a 14, numa escala de 0 a 20.

Uma vez que um dos objetivos desta dissertação é compreender em que medida os métodos de avaliação são capazes de explicar a variável média, será utilizado para este estudo a modelação de equações estruturais. Segundo Amorim et al., (2012) a modelação com equações estruturais (ou SEM – *Structural Equation Modeling*) consiste na utilização de técnicas multivariadas de análise de dados que combinam pontos da regressão múltipla e de análise fatorial para estimar simultaneamente uma série de relações de dependência.

Para construir um modelo de equações estruturais é necessária a construção prévia de um modelo teórico que permitirá determinar as múltiplas relações de dependência entre as variáveis do modelo, a saber: 1ª Etapa: Construção de um modelo teórico - consiste em expressar as hipóteses do investigador baseando-se nas relações de causa entre as variáveis selecionadas para a pesquisa com fundamentos teóricos; 2ª Etapa: Construção do diagrama de caminhos das relações definidas - expressão gráfica de causa e efeito desenvolvida no modelo teórico; 3ª Etapa: Conversão do diagrama de caminhos num modelo estrutural. A partir do modelo teórico são criadas as hipóteses a serem testadas na aplicação dos sistemas de equações estruturais; 4ª Etapa: Verificação da existência de limitações explanatórias do modelo para gerar soluções únicas; 5ª Etapa: Avaliação dos critérios de ajuste do modelo - etapa responsável pela avaliação da adequação e ajuste geral do modelo através do uso das medidas de adequação absolutas, medidas de ajuste incremental e medidas de ajustes de parcimônia; e 6ª Etapa: Interpretação e modificação do modelo - consiste na comparação dos resultados obtidos no desenvolvimento do modelo com os objetos e hipóteses estabelecidas no estudo a partir da teoria, comparando estatisticamente essas afirmações.

Dado que o passo seguinte à modelação de equações estruturais será prever como se comportariam as médias, será utilizada a metodologia KDD (*Knowledge Discovery in Database*). Pimentel e Omar (2006) realizaram um estudo cujo objetivo consistiu em apresentar um modelo para a aplicação de técnicas Data Mining, utilizando algoritmos de extração de padrões para a descoberta de conhecimento através dos dados obtidos nas avaliações de desempenho de um grupo de alunos. Nesse estudo, os autores também aplicaram as várias etapas do processo KDD, discriminadas na Tabela 3.1.

Tabela 3.1: Fases e Etapas do processo de KDD (Pimentel & Omar, 2006)

FASE	ETAPAS	OBJETIVO
Preparação	Seleção dos dados	Compreender o domínio da aplicação, estabelecer objetivos e definir a massa de dados inicial a ser utilizada
	Pré-processamento e Limpeza	Eliminar eventuais inconsistências e problemas com os dados e limpeza dos mesmos.
	Transformação dos dados	Transformar os dados para uma forma mais apropriada para a mineração, assumindo um formato mais estatístico e menos transacional

Extração de Padrões	Mineração dos Dados	Aplicar métodos (algoritmos) para a Extração de Padrões
	Avaliação dos Padrões	Analisar os resultados da etapa anterior a fim de identificar conhecimento nos padrões encontrados
Pré-Processamento	Apresentação do Conhecimento	Utilizar técnicas de visualização e representação do conhecimento para apresentar ao utilizador o conhecimento minerado

Neste estudo, os autores Pimentel e Omar (2006) utilizaram dados recolhidos em quatro sessões de avaliação, com 52 participantes que responderam a 20 problemas de Lógica de Programação. Estes problemas envolviam 12 conceitos diferentes e obtiveram um total de 3277 registos. As avaliações foram formuladas de forma a obter uma medida cognitiva e duas medidas metacognitivas (para as medidas metacognitivas os autores optaram por utilizar o *Knowledge Monitoring Accuracy* (KMA); e o *Knowledge Monitoring Bias* (KMB). Para a aplicação de técnicas de Data Mining os autores utilizaram, neste estudo, a ferramenta Weka (*Waikato Environment for Knowledge Analysis*). As técnicas de Data Mining utilizadas para o estudo foram a Classificação e a Associação.

Pimentel e Omar (2006) concluíram que o processo de avaliação continua a poder gerar um grande volume de dados, que necessita de ser tratado e interpretado de modo a fornecer informações mais precisas sobre o conhecimento atual do aluno. Os autores constataam, que o volume de dados se pode tornar maior se pensarmos na possibilidade de cruzar disciplinas e pré-requisitos. As técnicas de Data Mining aplicadas no estudo demonstraram ser viáveis e aplicáveis nos dados recolhidos através das avaliações. Os resultados obtidos pelos autores demonstraram as potencialidades destas técnicas, no entanto é necessário um aprofundamento nos testes e análises dos padrões encontrados no estudo.

3.2 Descrição da recolha e tratamento dos dados

Nesta fase foi necessário proceder à recolha, à devida filtragem e ao tratamento dos dados. A recolha de dados relacionados com as médias das Unidades Curriculares entre os anos letivos 2014-2015 e 2018-2019 foi conseguida através da ferramenta produzida por Miguel, Ramos, Martins e Costa (2020). A recolha de dados sobre os métodos de avaliação foi efetuada a partir do separador “Métodos de avaliação” das Fichas das Unidades Curriculares.

A Figura que se segue (Figura 3.1) representa o layout de pesquisa de Miguel et al. (2020). Neste separador é possível selecionar o tipo de informação pretendida através da execução de uma *query*.

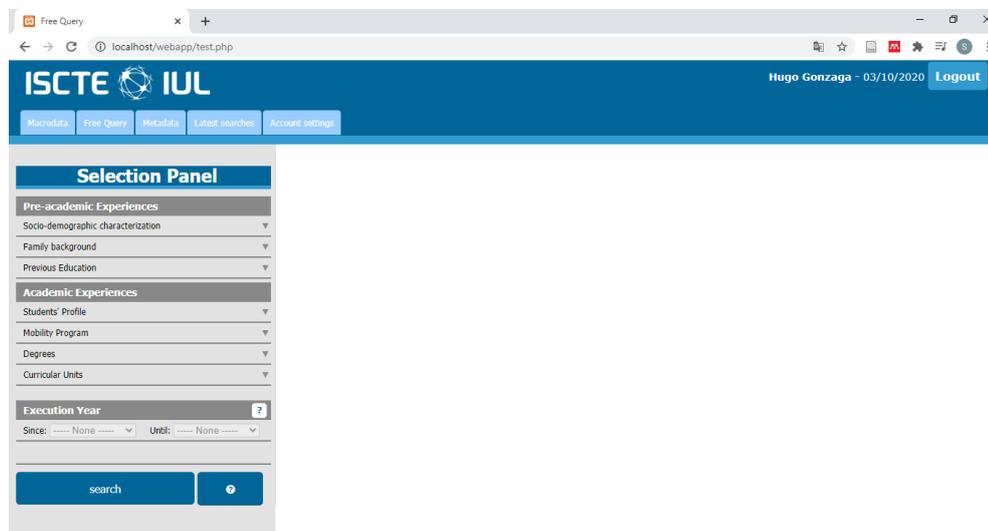


Figura 3.1: Framework da aplicação de Miguel et al. (2020)

A Figura 3.2 apresenta as opções selecionadas para o efeito do estudo:

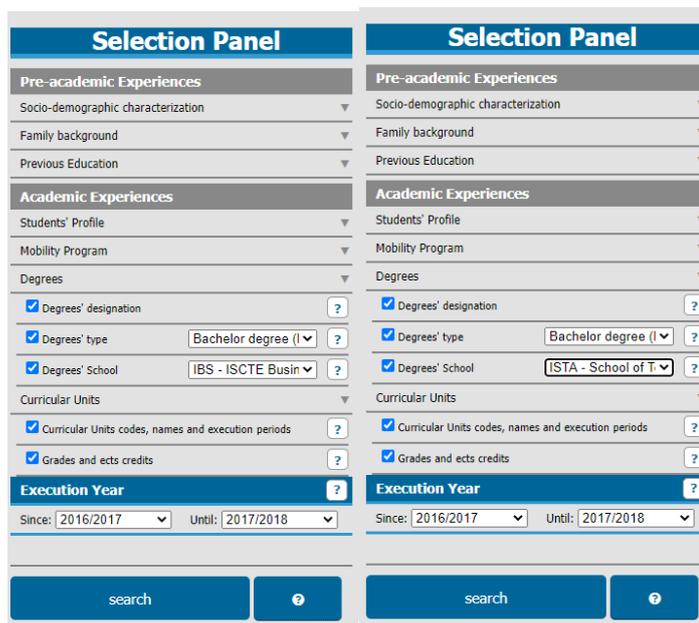


Figura 3.2: Layout de pesquisa da Aplicação de Miguel et al. (2020)

Foram selecionadas as opções em relação aos cursos “Designação do curso”, “Tipo de curso” e “Escola do curso” e em relação às unidades curriculares foram selecionados os dois campos, “Unidades curriculares códigos, nomes e períodos de execução” e “Notas e créditos ects” e ainda foi considerado o campo “Ano de execução”. Uma vez que era necessário extrair apenas as médias dos cursos pertencentes a duas escolas e que fossem apenas do tipo licenciatura. Em relação às Unidades Curriculares foram selecionadas e utilizadas para o estudo complementar

onde poderá ser possível perceber a que grupo pertencem as Unidades Curriculares com média mais alta.

Após a seleção das “*freequerys*” a ferramenta produziu um ficheiro que pode ser exportado para o Excel com as colunas selecionadas. Para o estudo apenas foram necessárias informações referentes ao nome das Unidades Curriculares e respetiva classificação média.

	curricularUnitCode	curricularUnitName	period	gradeValue	gradeScale
ifecture	L2102		2º Ano, 2º Semestre	NA	TYPE20
ifecture	L0787		3º Ano, 1º Semestre	NA	TYPE20
ifecture	L5096		2º Ano, 1º Semestre	RE	TYPE20
ifecture	L5106		3º Ano, 2º Semestre	10	TYPE20
ifecture	01766		1º Ano, 2º Semestre	14	TYPE20
ifecture	L5096		2º Ano, 1º Semestre	8	TYPE20
ifecture	L5096		2º Ano, 1º Semestre	13	TYPE20
ifecture	L5102		2º Ano, 2º Semestre	NA	TYPE20
ifecture	L0141		1º Ano, 1º Semestre	13	TYPE20
ifecture	L5106		3º Ano, 2º Semestre	10	TYPE20
ifecture	L5105		3º Ano, 1º Semestre	10	TYPE20
ifecture	01769		1º Ano, 2º Semestre	12	TYPE20
ifecture	L0779		3º Ano, 1º Semestre	NA	TYPE20
ifecture	L0743		3º Ano, 2º Semestre	RE	TYPE20
ifecture	L5098		3º Ano, 1º Semestre	NA	TYPE20
ifecture	L0731		3º Ano, 1º Semestre	RE	TYPE20
ifecture	L0726		2º Ano, 1º Semestre	NA	TYPE20
ifecture	L5099		3º Ano, 2º Semestre	NA	TYPE20
ifecture	L0731		3º Ano, 1º Semestre	13	TYPE20
ifecture	02864		1º Ano, 2º Semestre	NA	TYPE20
ifecture	L2102		2º Ano, 2º Semestre	NA	TYPE20
ifecture	L0131		1º Ano, 1º Semestre	NA	TYPE20
ifecture	L0131		1º Ano, 1º Semestre	NA	TYPE20

This query returned: 26074 cases

Figura 3.3: Resultado da pesquisa

De seguida, iniciou-se o processo de recolha manual das informações sobre os métodos de avaliação presentes em todas as FUCs (Fichas de Unidade Curricular) das UCs (Unidades Curriculares) pertencentes a duas escolas do ISCTE com cursos nas áreas de Tecnologias e Gestão. O processo de recolha de dados das FUCs demorou cerca de três semanas.

Foi utilizada a ferramenta Excel para trabalhar os dados, juntando os métodos de avaliação presentes nas FUCs e as respetivas classificações médias de todos os anos disponíveis na amostra. A Figura 3.4 representa a organização da base de dados deste estudo. As variáveis tidas em conta foram: Cadeira, AnoLetivo, Projeto, PesoProj, Mini-Testes, PesoMT, 1Teste, Peso1T, 2Testes, Peso2T, Exame, PesoE, Trabalho de Grupo, PesoTG, Trabalho Individual, PesoTI, Frequência, PesoF, Exercícios, PesoEx, Participação, PesoP, Laboratórios, PesoL Apresentação, PesoA e Media. Todas as variáveis “peso” foram retiradas das FUCs.

Figura 3.4: Organização dos dados (Excel)

Posteriormente os dados foram trabalhados utilizando o software SPSS Statistics (v.26) onde foi possível verificar se os dados estavam bem tratados, isto é, se não continham valores errados nem *missings values*. Na Tabela 3.2 é possível observar que todas as variáveis tinham valores válidos e que também não existiam *missing values*.

Tabela 3.2: Frequências

		Estatísticas					
		Cadeira	Ano Letivo	PesoProj	PesoMT	PesoTI	PesoF
n	Válido	796	796	796	796	796	796
	Missing	0	0	0	0	0	0
		PesoEx	Peso2T	PesoE	PesoTG	PesoL	PesoA
n	Válido	796	796	796	796	796	796
	Missing	0	0	0	0	0	0
		PesoP	Peso1T				
n	Válido	796	796				
	Missing	0	0				

4. Resultados

4.1 Análise de dados

Após os dados serem recolhidos e filtrados, foi iniciada a análise descritiva das variáveis.

Para perceber melhor o comportamento das variáveis a utilizar no estudo das equações estruturais apresenta-se uma análise mais aprofundada às variáveis “peso” onde será possível perceber se existiam ou não relações entre elas. As variáveis escolhidas foram as variáveis numéricas “peso” uma vez que as variáveis binárias não teriam tanto interesse a nível estatístico por não conterem tanta informação.

Sendo as hipóteses estatísticas:

H_0 : Não existe relação entre as variáveis “peso” e a média.

H_a : Existe relação entre as variáveis “peso” e a média.

Na Tabela 4.1 confirma-se, através do teste Kolmogorov-Smirnov (K-S) de uma amostra, que nenhuma das variáveis segue a distribuição normal, uma vez que, em todos os casos o valor do Sig é inferior a 0,05, no entanto como o teste K-S é sensível a grandes amostras apresenta-se a análise da Assimetria e Curtose.

Tabela 4.1: Teste de Kolmogorov-Smirnov de Uma Amostra

		Media	PesoA	PesoProj	PesoMT	Peso1T	Peso2T
n		792	796	796	796	796	796
Parâmetros normais	Média	12,98	2,2	2,71	3,44	16,17	0,81
	Desvio padrão	1,63	7,3	12,08	10,41	19,09	5,85
Diferenças mais extremas	Absoluto	0,03	0,51	0,53	0,49	0,35	0,53
	Positivo	0,02	0,51	0,53	0,49	0,35	0,53
	Negativo	-0,03	-0,38	-0,41	-0,37	-0,2	-0,45
TestStatistic		0,03	0,51	0,53	0,49	0,35	0,53
valor-p		,027	<0,001	<0,001	<0,001	<0,001	<0,001

		PesoE	PesoTG	PesoF	PesoEx	PesoP	PesoL
n		796	796	796	796	796	796
Parâmetros normais	Média	0	17,01	46,1	2,51	3,3	2,45
	Desvio padrão	,00	18,33	15,9	8,12	6,33	8,69
Diferenças mais extremas	Absoluto		0,31	0,22	0,5	0,44	0,53
	Positivo		0,31	0,22	0,5	0,44	0,53
	Negativo		-0,18	-0,2	-0,38	-0,3	-0,39
TestStatistic			0,31	0,22	0,5	0,44	0,53
valor-p			<0,001	<0,001	<0,001	<0,001	<0,001

Com o auxílio da ferramenta Excel foi possível calcular o valor da distribuição através do valor da assimetria e da curtose, este valor foi calculado dividindo o valor do coeficiente de assimetria (Skewness) pelo valor do erro padrão da assimetria e dividindo o coeficiente de curtose (Kurtosis) pelo valor do erro padrão da curtose. Para que as variáveis sigam a distribuição normal é necessário que os dois valores estejam compreendidos entre -2 e 2.

Na Tabela 4.2 é possível observar que nas colunas “Valor da distribuição” que não existe nenhum valor compreendido entre -2 e 2. Por esse motivo, conclui-se que nenhuma das variáveis segue a distribuição normal.

Uma vez que nenhuma das variáveis segue a distribuição normal foi utilizada a correlação de Spearman para a análise de correlações uma vez que este índice é o indicado para o tamanho da amostra ($n > 30$) e para variáveis que não seguem a distribuição normal (Marôco, 2014 b).

Tabela 4.2: Assimetria e Curtose

	Skewness			Kurtosis		
	Estatísticas	Erro padrão	Valor da distribuição	Estatísticas	Erro padrão	Valor da distribuição
PesoProj	5,46	0,09	62,99	33,49	0,17	193,48
PesoMT	3,71	0,09	42,86	14,45	0,17	83,48
Peso1T	0,52	0,09	6,04	-1,43	0,17	-8,23
Peso2T	9,85	0,09	113,70	107,80	0,17	622,76
PesoE		0,09	0,00		0,17	0,00
PesoTG	0,49	0,09	5,63	-1,16	0,17	-6,72
PesoTI	5,12	0,09	59,07	33,39	0,17	192,90
PesoF	-0,65	0,09	-7,53	2,92	0,17	16,88
PesoEx	3,82	0,09	44,03	15,06	0,17	87,03
PesoP	2,06	0,09	23,72	4,22	0,17	24,41
PesoL	3,48	0,09	40,19	10,79	0,17	62,33
PesoA	3,72	0,09	42,92	13,83	0,17	79,90
Media	-0,48	0,09	-5,54	0,56	0,17	3,25

Para o estudo das correlações a variável **PesoE**, peso do exame, foi retirada, uma vez que esta não apresentava valores.

Na Tabela 4.3 observam-se algumas correlações que merecem particular atenção.

Em primeiro lugar, verifica-se que o peso que afeta mais a variável média é o peso do trabalho de grupo (**PesoTG**) que tem valor de 0,360. Embora esta seja uma correlação baixa, corresponde a uma relação linear ascendente, o que significa que afeta positivamente a variável média e conseqüentemente aumenta o seu valor. Em seguida verificou-se que o peso das apresentações (**PesoA**) com 0,210, o peso dos trabalhos individuais (**PesoTI**) com valor de 0,202, o peso do projeto (**PesoProj**) com 0,169 e o peso da participação (**PesoP**) com valor de 0,136, têm também coeficientes de correlação positivos, e conseqüentemente também afetam a variável média positivamente embora tenham uma relação linear fraca.

Relativamente às variáveis que afetam a média negativamente, isto é, que contribuem para a diminuição da média, observa-se que o peso de um teste (**Peso1T**) é o que mais afeta a variável com o valor de -0,290 de seguida o peso dos laboratórios (**PesoL**) com -0,270, o peso da frequência (**PesoF**) com o valor de -0,137, o peso dos mini testes (**PesoMT**) com -0,123, o peso de dois testes (**Peso2T**) com -0,056 e finalmente o peso dos exercícios (**PesoEx**) com -0,014, sendo este o que menos afeta a variável. Importa referir que todos os valores dos coeficientes de correlação negativos são de fraca relação linear.

É ainda possível, concluir-se através do valor-p do Ró de Spearman's, que as variáveis Peso2T e PesoEx, que apresentam valores de valor-p superiores a 0,005, por isso existem evidências estatísticas para se aceitar H_0 , não existindo assim relação entre estas variáveis e a variável média. No entanto, para as restantes variáveis uma vez que, o valor -p é inferior a 0,005, existem evidências estatísticas para se aceitar H_a , existindo assim relação entre essas variáveis “peso” e a variável média.

Tabela 4.3: Correlação

			PesoP	PesoA	PesoProj	PesoMT	Peso1T	Peso2T
Spearman's rho	Media	Coefficiente de Correlação	,136**	,210**	,169**	-,123**	-,290**	-0,056
		valor-p	<0,001	<0,001	<0,001	0,001	<0,001	0,118
		n	792	792	792	792	792	792
			PesoTG	PesoTI	PesoL	PesoF	PesoEx	
Spearman's rho	Media	Coefficiente de Correlação	,360**	,202**	-,270**	-,137**	-0,014	
		valor-p	<0,001	<0,001	<0,001	<0,001	0,704	
		n	792	792	792	792	792	

4.2 Modelação de equações estruturais

No subcapítulo anterior observou-se que existiam correlações entre as variáveis “peso” e a variável média. Foi ainda possível verificar-se, que para a maioria das variáveis “peso” existia uma relação com a variável média e pelos motivos apresentados justifica-se o estudo da forma em que essas variáveis “peso” são capazes de explicar a variável média.

Este estudo teve por base os procedimentos de Marôco (2014 a) para a modelação de equações estruturais. Neste subcapítulo, foram consideradas as seguintes variáveis:

- PesoL- Peso laboratórios
- PesoProj – Peso projeto
- PesoMT – Peso mini testes
- Peso1T – Peso de um teste intermédio
- Peso2T – Peso de 2 testes intermédios
- PesoTG – Peso trabalho de grupo
- PesoA – Peso apresentação
- PesoF – Peso frequência
- PesoTI – Peso trabalho individual
- PesoP – Peso participação
- PesoEx – Peso exercícios
- Média

Foi descartada a variável PesoE, peso do exame, uma vez que após a análise de frequências para confirmação e consolidação dos dados não foram observadas ocorrências do PesoE, como se pode observar na tabela abaixo (Tabela 4.4).

Tabela 4.4: Frequências PesoE

PesoE					
		Frequência	Percentagem	% valida	% Acumulativa
Válido	0	796	100,0	100,0	100,0

Nos próximos subcapítulos apresentam-se três estudos baseados nas questões de investigação. Tendo o primeiro estudo como base todos os valores possíveis para a média, isto é valores compreendidos entre 0 e 20, o segundo estudo valores de média iguais ou superiores a 10 e finalmente o terceiro estudo para valores de média iguais ou superiores a 14.

4.2.1 Estudo 1

Começando pela análise da base de dados completa, isto é, com todos os valores para a variável média e com todas as variáveis “peso” recolhidas apresenta-se a Tabela 4.5.

Na Tabela 4.5 a coluna importante para o estudo é a P. Para que o modelo seja o melhor possível todos os valores P (Pvalue) têm de ter o valor “****”, isto significa que $P < 0,001$. Conclui-se assim que o modelo não é de todo o mais indicado.

Tabela 4.5: Pesos Regressão Modelo 1 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,020	,014	-1,454	,146
Media	<---	PesoProj	,046	,013	3,535	***
Media	<---	PesoMT	,021	,013	1,609	,108
Media	<---	Peso1T	,011	,013	,892	,372
Media	<---	Peso2T	,034	,017	2,046	,041
Media	<---	PesoTG	,043	,013	3,393	***
Media	<---	PesoA	,050	,014	3,524	***
Media	<---	PesoF	,010	,013	,771	,440
Media	<---	PesoTI	,034	,013	2,597	,009
Media	<---	PesoP	,032	,015	2,175	,030
Media	<---	PesoEx	,020	,014	1,464	,143

Na Figura 4.1 observa-se que o valor da média é explicado em 29%, isto significa que todas as variáveis juntas explicam 29% da variação observada da variável média. Verifica-se ainda que a variável PesoL contribui para a média de forma inversa. Isto significa que a sua existência prejudica a média dos alunos.

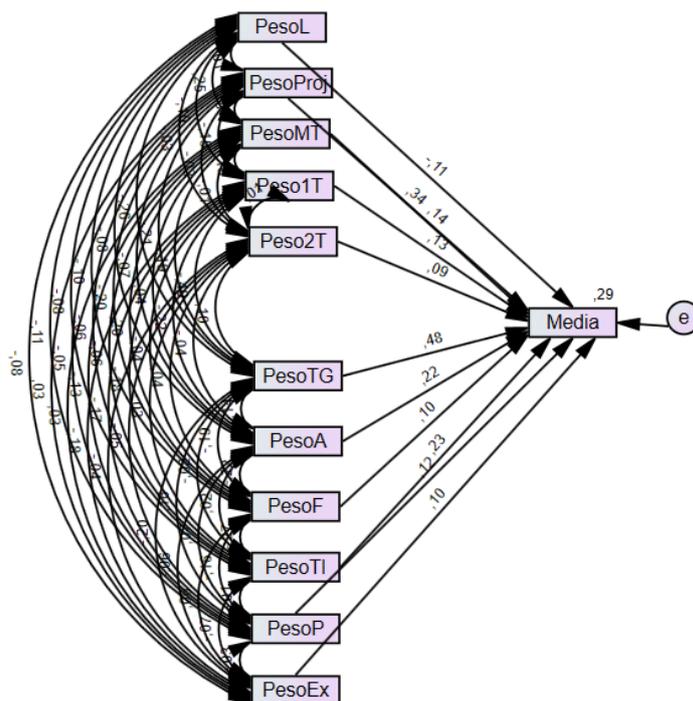


Figura 4.1: Modelo 1 Estudo 1

Tabela 4.6: Coeficientes e Estatísticas de Colinearidade

Modelo 1	Coeficientes(a)							
	Coeficientes não Normalizado		Coeficientes Normalizado		t	Sig.	Estatísticas De Colinearidade	
	B	Erro padrão	Beta				Tolerância	VIF
(Constante)	11,065	1,260			8,785	0,000		
PesoProj	0,046	0,013	0,344		3,511	0,000	0,095	10,476
PesoMT	0,021	0,013	0,138		1,598	0,110	0,124	8,088
Peso1T	0,011	0,013	0,133		0,886	0,376	0,040	24,756
Peso2T	0,034	0,017	0,092		2,032	0,043	0,446	2,245
PesoTG	0,043	0,013	0,484		3,369	0,001	0,044	22,506
PesoTI	0,034	0,013	0,233		2,578	0,010	0,112	8,933
PesoF	0,010	0,013	0,096		0,766	0,444	0,059	16,967
PesoEx	0,020	0,014	0,101		1,453	0,147	0,189	5,301
PesoP	0,032	0,015	0,124		2,160	0,031	0,280	3,576
PesoL	-0,020	0,014	-0,107		-1,444	0,149	0,168	5,965
PesoA	0,050	0,014	0,224		3,499	0,000	0,223	4,483

a. Variável dependente: Media

De forma a ajudar a elaboração de equações estruturais é importante observar-se a tabela de Coeficientes e Estatísticas de Colinearidade, onde a variável dependente é a variável média. No entanto, para esta análise é importante observar a coluna VIF da Tabela 4.6 que representa o valor da multicolinearidade, uma vez que um grau elevado de multicolinearidade diminui a precisão com que se estimam os parâmetros dos modelos. Por isso passam a existir grandes oscilações nas estimativas dos parâmetros do modelo o que dificulta a interpretação dos coeficientes estimados. (Pestana & Gageiro, 2014). Pode-se ainda fazer a análise à regressão linear utilizando a opção STEPWISE que indica as variáveis dos melhores modelos de regressão linear.

Na Tabela 4.7 é possível observar os diferentes modelos produzidos pelo SPSS V26 e o respetivo valor de R2 ajustado. O modelo com o maior R2, isto é, com a maior percentagem de variação na resposta que é explicada é o modelo 8.

Tabela 4.7: Sumário dos Modelos

Sumario Modelos										
Modelo	R	R Quadrado	R Quadrado Ajustado	Estimativa Erro Padrão	Estatísticas Aletradas					Durbin-Watson
					R Quadrado Alterado	F Alterado	df1	df2	Sig. F Alterado	
1	,357a	0,127	0,126	1,526	0,127	115,004	1	789	0,000	
2	,428b	0,183	0,181	1,477	0,056	53,693	1	788	0,000	
3	,467c	0,218	0,215	1,446	0,035	35,003	1	787	0,000	
4	,498d	0,248	0,244	1,419	0,030	31,689	1	786	0,000	
5	,520e	0,270	0,265	1,399	0,022	23,717	1	785	0,000	
6	,525f	0,275	0,270	1,395	0,005	5,643	1	784	0,018	
7	,529g	0,280	0,273	1,391	0,004	4,675	1	783	0,031	
8	,532h	0,283	0,276	1,389	0,004	3,919	1	782	0,048	0,985

a. Preditores: (Constante), PesoTG; b. Preditores: (Constante), PesoTG, PesoProj; c. Preditores: (Constante), PesoTG, PesoProj, PesoA; d. Preditores: (Constante), PesoTG, PesoProj, PesoA, PesoL; e. Preditores: (Constante), PesoTG, PesoProj, PesoA, PesoL, PesoTI; f. Preditores: (Constante), PesoTG, PesoProj, PesoA, PesoL, PesoTI, PesoP; g. Preditores: (Constante), PesoTG, PesoProj, PesoA, PesoL, PesoTI, PesoP, PesoMT; h. Preditores: (Constante), PesoTG, PesoProj, PesoA, PesoL, PesoTI, PesoP, PesoMT, Peso2T; i. Variavel dependente: Media

Na Tabela do Anexo A, são apresentados os 8 modelos produzidos pela opção STEPWISE, no entanto, para estudo começa-se por considerar apenas o ultimo modelo (modelo 8 Tabela 4.8) de regressão linear produzido, uma vez que contém mais variáveis e à partida será o que contém uma maior variação observada da variável média.

Tabela 4.8: Coeficientes Stepwise Estudo 1

Modelo	Coeficientes Não Normalizado	Coeficientes		t	Sig.	Estatística de Colinearidade		
		B	Std. Error			Beta	Tolerância	VIF
8	(Constante)	12,150	0,087					
	PesoTG	0,031	0,003	0,352	10,634	0,000	0,834	1,198
	PesoProj	0,036	0,004	0,264	8,427	0,000	0,933	1,072
	PesoA	0,039	0,007	0,175	5,710	0,000	0,974	1,027
	PesoL	-0,032	0,006	-0,171	-5,275	0,000	0,873	1,146
	PesoTI	0,023	0,004	0,158	5,149	0,000	0,979	1,021
	PesoP	0,021	0,008	0,083	2,670	0,008	0,957	1,045
	PesoMT	0,011	0,005	0,071	2,216	0,027	0,899	1,113
	Peso2T	0,022	0,011	0,061	1,980	0,048	0,977	1,023

A Figura 4.2 representa a equação estrutural do modelo 8 de regressão linear produzida pelo programa SPSS V26, consegue-se observar que a variação observada da variável média é de 28% diminuindo assim apenas em 1%. No entanto na Tabela 4.9 verifica-se que as variáveis PesoMT e Peso2T não contêm um valor de PValue significativo, isto porque os valores são muito superiores a 0,001, por esse motivo ambas serão removidas para a elaboração do próximo modelo.

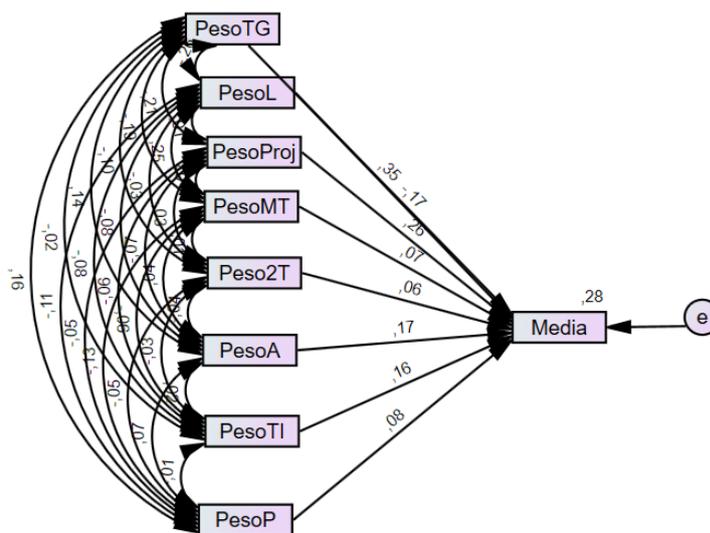


Figura 4.2: Modelo 2 Estudo 1

Embora aconteça o mesmo no PesoP, isto é o valor de P é superior a 0,001 não é tão elevado como as variáveis referidas anteriormente e por esse motivo não será removida para já.

Tabela 4.9: Pesos Regressão Modelo 2 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,032	,006	-5,302	***
Media	<---	PesoProj	,036	,004	8,470	***
Media	<---	PesoMT	,011	,005	2,227	,026
Media	<---	Peso2T	,022	,011	1,990	,047
Media	<---	PesoA	,039	,007	5,739	***
Media	<---	PesoTI	,023	,004	5,175	***
Media	<---	PesoP	,021	,008	2,684	,007
Media	<---	PesoTG	,031	,003	10,688	***

Após a remoção das variáveis Peso2T e PesoMT observa-se na Figura 4.3 que variação observada da variável média é de 27% embora tenha voltado a diminuir 1% a Tabela 4.10 apresenta apenas um valor de P não significativo, sendo este o do PesoP.

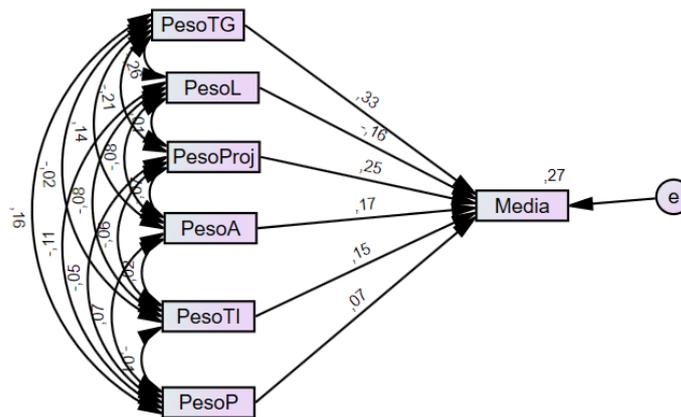


Figura 4.3: Modelo 3 Estudo 1

Tabela 4.10: Pesos Regressão Modelo 3 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,030	,006	-5,088	***
Media	<---	PesoProj	,034	,004	8,132	***
Media	<---	PesoA	,039	,007	5,644	***
Media	<---	PesoTI	,022	,004	4,941	***
Media	<---	PesoP	,019	,008	2,385	,017
Media	<---	PesoTG	,030	,003	10,268	***

O próximo modelo apresentado conta com a remoção da variável PesoP, uma vez que na tabela peso da regressão percebe-se que o PValue aumentou de 0,007 para 0,017 ficando assim menos significante para o estudo.

Assim, o modelo que se segue cumpre todos os requisitos para que seja considerado o melhor modelo uma vez que é capaz de explicar 27% da variável média e contém todos os valores de PValue significativos (Ver Tabela 4.11).

Na Figura 4.4 nota-se que a maioria das variáveis contribuem positivamente para a média, isto é, quanto maior peso for atribuído mais alta será a média, à exceção da variável PesoL, peso dos laboratórios, que contém uma relação inversa com a variável média, isto significa que quanto maior for o peso atribuído aos laboratórios menor será o valor da média.

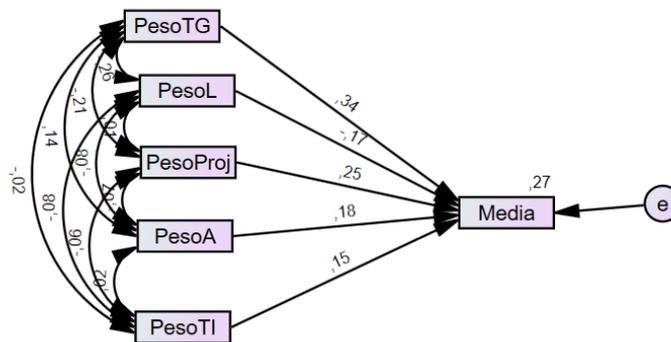


Figura 4.4: Modelo 4 Estudo 1

Tabela 4.11: Pesos Regressão Modelo 4 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,031	,006	-5,247	***
Media	<---	PesoProj	,034	,004	8,046	***
Media	<---	PesoA	,039	,007	5,740	***
Media	<---	PesoTI	,022	,005	4,886	***
Media	<---	PesoTG	,031	,003	10,590	***

Pode-se ainda fazer a remoção das variáveis através da observação da coluna P das diversas tabelas de pesos da regressão obtidas através do programa AMOS V26.

Começando pela análise da primeira tabela, Tabela 4.12, em que o modelo é exatamente o mesmo modelo inicial, isto é, contém todas as variáveis “peso”. Observa-se que existe um número elevado de variáveis não significativas, uma vez que estas contêm um valor bastante elevado de PValue, como por exemplo PesoF, Peso1T, PesoL e PesoMT.

Tabela 4.12: Pesos Regressão Modelo 5 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,020	,014	-1,454	,146
Media	<---	PesoProj	,046	,013	3,535	***
Media	<---	PesoMT	,021	,013	1,609	,108
Media	<---	Peso1T	,011	,013	,892	,372
Media	<---	Peso2T	,034	,017	2,046	,041
Media	<---	PesoTG	,043	,013	3,393	***
Media	<---	PesoA	,050	,014	3,524	***
Media	<---	PesoF	,010	,013	,771	,440
Media	<---	PesoTI	,034	,013	2,597	,009
Media	<---	PesoP	,032	,015	2,175	,030
Media	<---	PesoEx	,020	,014	1,464	,143

Após a remoção todas as variáveis com o valor de P elevado, e ficando apenas com as variáveis PesoProj, Peso2T, PesoTG, PesoA, PesoTI e PesoP.

Obtém-se o seguinte modelo (Figura 4.5):

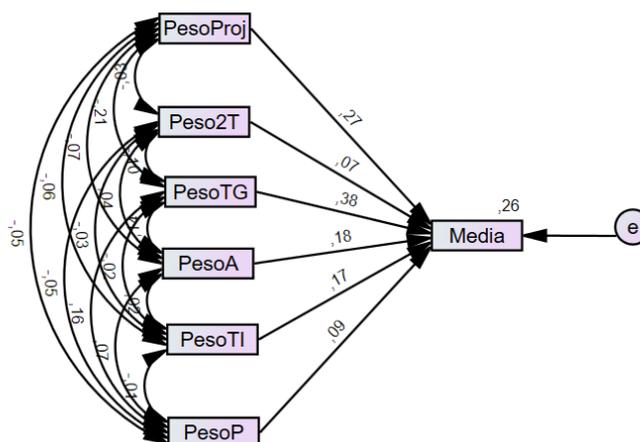


Figura 4.5: Modelo 5 Estudo 1

Observa-se que desta forma a variação explicada da variável média é inferior à estudada anteriormente com o auxílio do método de STEPWISE para a regressão linear. No entanto ao longo das próximas análises será apresentado o melhor modelo utilizando apenas a análise das tabelas dos pesos da regressão.

Tabela 4.13: Pesos Regressão Modelo 6 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoProj	,036	,004	8,449	***
Media	<---	Peso2T	,026	,011	2,266	,023
Media	<---	PesoTG	,034	,003	11,956	***
Media	<---	PesoA	,041	,007	5,863	***
Media	<---	PesoTI	,025	,005	5,462	***
Media	<---	PesoP	,022	,008	2,794	,005

Observando a Tabela 4.13 verifica-se que o próximo passo será remover as variáveis Peso2T e PesoP, embora esta última também tenha um grau de significância que poderia ser considerado no modelo. Na Figura 4.6, a variação explicada é de 24%, menos 3% que no método anterior. No entanto todas as variáveis “Peso” influenciam a média de forma positiva isto porque todas têm um efeito direto positivo em relação à média.

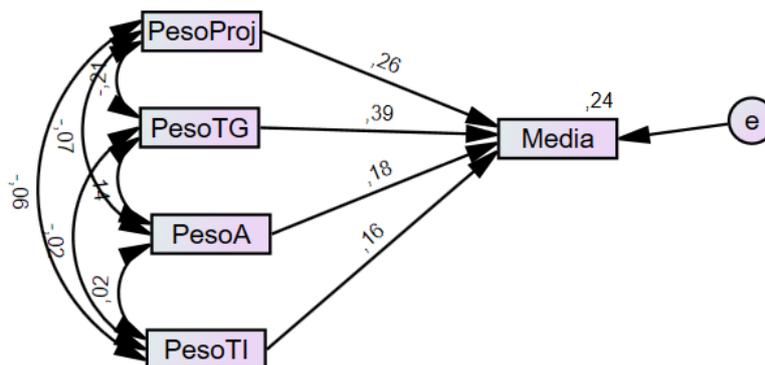


Figura 4.6: Modelo 6 Estudo 1

Ficando assim a Tabela 4.14 referente aos pesos da regressão com os seguintes valores de P:

Tabela 4.14: Pesos Regressão Modelo 7 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoProj	,035	,004	8,220	***
Media	<---	PesoTG	,035	,003	12,196	***
Media	<---	PesoA	,041	,007	5,914	***
Media	<---	PesoTI	,024	,005	5,308	***

Uma vez que seria interessante ver o impacto que a variável PesoP teria no modelo, dado que esta tinha anteriormente um valor de P igual a 0,005, determinou-se um modelo onde conste essa mesma variável.

Verifica-se na Figura 4.7 que o impacto dessa variável é de 1% uma vez que a variação explicada passou de 24% para 25%. Embora esta apresente também um efeito direto positivo na variável média que é apenas de 0,08.

No entanto na Tabela 4.15 observa-se que o valor de P passou de 0,005 para 0,007.

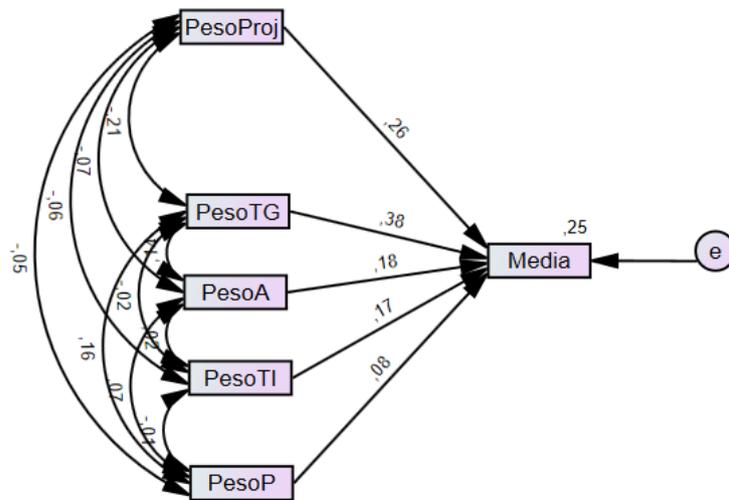


Figura 4.7: Modelo 7 Estudo 1

Tabela 4.15: Pesos Regressão Modelo 8 Estudo 1

			Estimativa	S.E.	C.R.	P
Media	<---	PesoProj	,035	,004	8,312	***
Media	<---	PesoTG	,034	,003	11,746	***
Media	<---	PesoA	,040	,007	5,799	***
Media	<---	PesoTI	,024	,005	5,359	***
Media	<---	PesoP	,022	,008	2,700	,007

Conclui-se então que o modelo para esta forma de seleção é constituído pelas variáveis PesoProj, PesoTG, PesoA e PesoTI.

4.2.2 Estudo 2

Na tabela seguinte (Tabela 4.16) verifica-se que o número de registos é inferior ao anterior uma vez que esta base de dados apenas contém valores de media positivos, isto é, valores iguais ou superiores a 10 valores.

Tabela 4.16: Frequências (Valores média >= 10)

		Estatísticas						
		PesoProj	PesoMT	Peso1T	Peso2T	PesoE	PesoTG	PesoTI
n	Válido	757	757	757	757	757	757	757
	Missing	0	0	0	0	0	0	0
		PesoF	PesoEx	PesoP	PesoL	PesoA	Media	
n	Válido	757	757	757	757	757	757	
	Missing	0	0	0	0	0	0	

Uma vez que anteriormente se concluiu que o método com uma maior variação explicada da variável média, foi utilizando a regressão linear com o auxílio do SPSS V26 e utilizando a opção de STEPWISE, para o estudo que se segue foi utilizada apenas essa metodologia.

Na Tabela 4.17 e na Figura 4.8 correspondentes ao primeiro modelo, são utilizadas todas as variáveis “peso”.

Observa-se que a variação explicada é de 27% e que o PesoL é a única variável que contribui negativamente para a média (Ver Figura 4.8).

Tabela 4.17: Pesos Regressão Modelo 1 Estudo 2

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,003	,012	-,268	,789
Media	<---	PesoProj	,046	,011	4,000	***
Media	<---	PesoMT	,019	,012	1,661	,097
Media	<---	Peso1T	,015	,011	1,379	,168
Media	<---	Peso2T	,033	,015	2,263	,024
Media	<---	PesoTG	,044	,011	3,950	***
Media	<---	PesoA	,051	,012	4,077	***
Media	<---	PesoF	,015	,011	1,320	,187
Media	<---	PesoTI	,036	,012	3,101	,002
Media	<---	PesoP	,031	,013	2,424	,015
Media	<---	PesoEx	,031	,012	2,525	,012

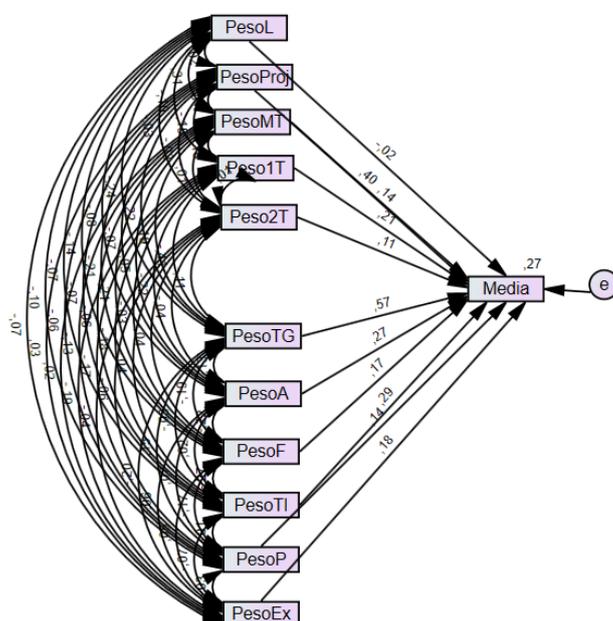


Figura 4.8: Modelo 1 Estudo 2

Em seguida apresenta-se o modelo de equação estrutural referente ao modelo de regressão linear 7 segundo STEPWISE, apresentado na Tabela 4.18. (Ver Tabela do Anexo B os restantes modelos).

Tabela 4.18: Coeficientes Stepwise Estudo 2

		Coeficientes ^(a)						
Modelo		Coeficientes Não Normalizado		Coeficientes Normalizados	t	Sig.	Estatística de Colinearidade	
		B	Erro Padrão	Beta			Tolerância	VIF
7	(Constante)	12,38	0,08		157,03	0,00		
	PesoTG	0,03	0,00	0,36	10,58	0,00	0,83	1,21
	PesoProj	0,03	0,00	0,26	8,17	0,00	0,94	1,06
	PesoA	0,04	0,01	0,19	5,89	0,00	0,97	1,03
	PesoTI	0,02	0,00	0,16	5,14	0,00	0,98	1,02
	PesoL	-0,02	0,01	-0,10	-3,00	0,00	0,91	1,10
	PesoEx	0,02	0,01	0,09	2,67	0,01	0,94	1,07
	PesoP	0,02	0,01	0,07	2,15	0,03	0,97	1,03

a. Variável Dependente: Média

Observando a Figura 4.9 é possível verificar que a variação explicada da variável média é de 26% e que o peso dos laboratórios contribui negativamente para a variável média. Relativamente à tabela dos pesos da regressão linear (Tabela 4.19) apresentados pelo sistema AMOS V26, é possível concluir que o valor P não é significativo para as variáveis PesoP e PesoEx.

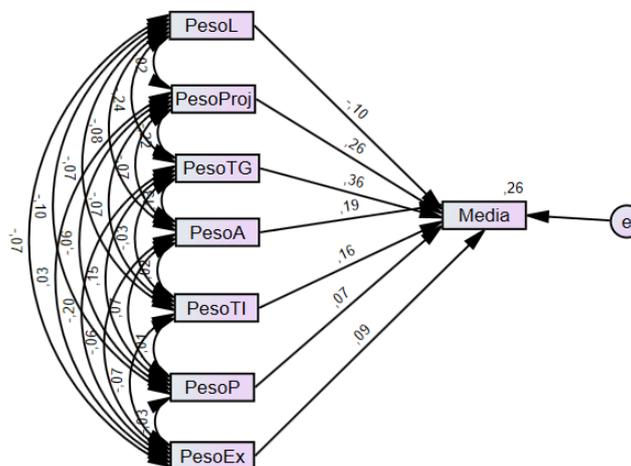


Figura 4.9: Modelo 2 Estudo2

Tabela 4.19: Pesos Regressão Modelo 2 Estudo2

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,017	,006	-3,014	,003
Media	<---	PesoProj	,030	,004	8,207	***
Media	<---	PesoTG	,028	,003	10,630	***
Media	<---	PesoA	,035	,006	5,915	***
Media	<---	PesoTI	,020	,004	5,161	***
Media	<---	PesoP	,015	,007	2,161	,031
Media	<---	PesoEx	,015	,006	2,687	,007

Eliminando essas variáveis fica-se com o seguinte modelo (Figura 4.10):

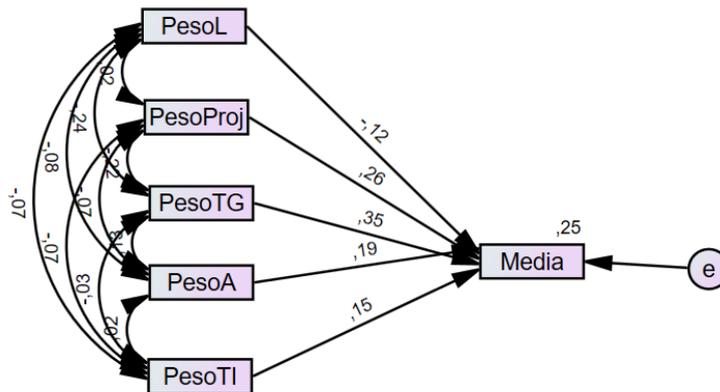


Figura 4.10: Modelo 3 Estudo 2

A Tabela 4.20 apresenta todos os valores de P significativos, permitindo concluir que o modelo apresentado é exatamente igual ao modelo de equações estruturais apresentado para a base de dados completa (com todos os valores de média existentes na amostra), embora a variação explicada seja inferior em 3%. Observa-se ainda que o PesoL continua a ser a única variável que tem um efeito negativo no valor da média.

Tabela 4.20: Pesos Regressão Modelo 3 Estudo2

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,020	,006	-3,535	***
Media	<---	PesoProj	,030	,004	8,013	***
Media	<---	PesoTG	,027	,003	10,525	***
Media	<---	PesoA	,035	,006	5,874	***
Media	<---	PesoTI	,019	,004	4,853	***

4.2.3 Estudo 3

Para a base de dados com valores de média iguais ou superiores a 14, o tamanho da amostra é apenas de 238. Para perceber melhor os dados começa-se por observar a tabela Coeficientes (Tabela 4.21) onde é possível perceber o valor de VIF de algumas variáveis. No entanto, o valor da variável o PesoTG (peso trabalho de grupo) será excluído da análise uma vez que este não está presente na Tabela 4.21 aparecendo numa tabela separada chamada “variáveis excluídas” como é possível observar na Tabela 4.22.

Tabela 4.21: Coeficientes Estudo 3

Modelo		Coeficiente ^(a)		t	Sig.	Estatísticas de Colinearidade		
		Coeficiente Não Normalizado	Coeficiente Normalizado			Tolerância	VIF	
		B	Erro Padrão					
1	(Constante)	15,224	0,197	77,235	0,000			
	PesoProj	0,004	0,003	0,091	1,279	0,202	0,701	1,427
	PesoMT	4,660E-05	0,006	0,001	0,008	0,993	0,746	1,340
	Peso1T	0,002	0,003	0,063	0,910	0,364	0,739	1,353
	Peso2T	-0,012	0,009	-0,083	-1,383	0,168	0,978	1,022
	PesoTI	-0,007	0,003	-0,157	-2,205	0,028	0,695	1,439
	PesoF	-0,012	0,003	-0,323	-3,675	0,000	0,455	2,199
	PesoEx	0,001	0,006	0,007	0,112	0,911	0,922	1,085
	PesoP	0,011	0,007	0,118	1,696	0,091	0,725	1,380
	PesoL	-0,012	0,008	-0,098	-1,616	0,108	0,961	1,040
	PesoA	0,007	0,005	0,110	1,492	0,137	0,651	1,536

a. Variável dependente: Media

Tabela 4.22: Variáveis Excluídas

Modelo		Variável excluída a						
		Beta In	t	Sig.	Correlação Parcial	Estatística de colinearidade		
						Tolerância	VIF	Tolerância Mínima
1	PesoTG	.b				0,000		0,000

a. Variável dependente: Media

b. Preditores no Modelo: (Constante), PesoA, PesoMT, PesoTI, Peso2T, PesoL, PesoEx, PesoP, PesoProj, Peso1T, PesoF

O seguinte modelo apresentado corresponde à utilização de todas as variáveis à exceção da variável PesoTG.

Através da Figura 4.11, observa-se que o modelo inicial é capaz de explicar apenas 20% e na Tabela 4.23 é possível observar que muitas das variáveis apresentam valores de P muito elevados.

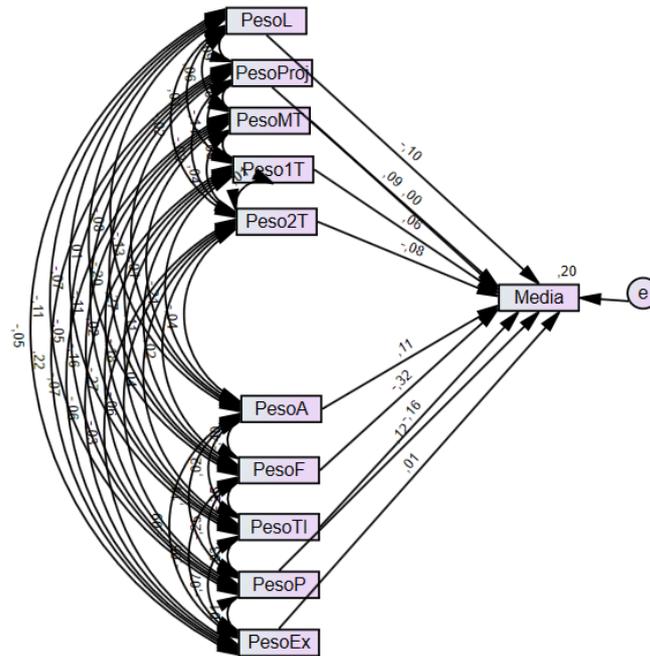


Figura 4.11: Modelo 1 Estudo 3

Tabela 4.23: Pesos Regressão Modelo 1 Estudo 3

			Estimativa	S.E.	C.R.	P
Media	<---	PesoL	-,012	,007	-1,651	,099
Media	<---	PesoProj	,004	,003	1,307	,191
Media	<---	PesoMT	,000	,005	,009	,993
Media	<---	Peso1T	,002	,003	,930	,353
Media	<---	Peso2T	-,012	,008	-1,413	,158
Media	<---	PesoA	,007	,004	1,524	,127
Media	<---	PesoF	-,012	,003	-3,755	***
Media	<---	PesoTI	-,007	,003	-2,253	,024
Media	<---	PesoP	,011	,007	1,733	,083
Media	<---	PesoEx	,001	,006	,114	,909

Analisando em seguida a Tabela 4.24 com os modelos de STEPWISE observa-se que este apresenta apenas dois modelos possíveis para esta base de dados.

Tabela 4.24: Coeficientes Stepwise Estudo 3

Modelo		Coeficiente a							
		Coeficiente Não Normalizado		Coeficiente Normalizado		t	Sig.	Estatísticas de Colinearidade	
		B	Erro Padrão	Beta	Tolerância			VIF	
1	(Constante)	15,332	0,101			152,326	0,000		
	PesoF	-0,013	0,002	-0,353		-5,789	0,000	1,000	1,000
2	(Constante)	15,482	0,110			140,639	0,000		
	PesoF	-0,015	0,002	-0,421		-6,600	0,000	0,880	1,136
	PesoTI	-0,009	0,003	-0,197		-3,094	0,002	0,880	1,136

a. Variável dependente: Média

Obtém-se então o seguinte modelo (Figura 4.12):

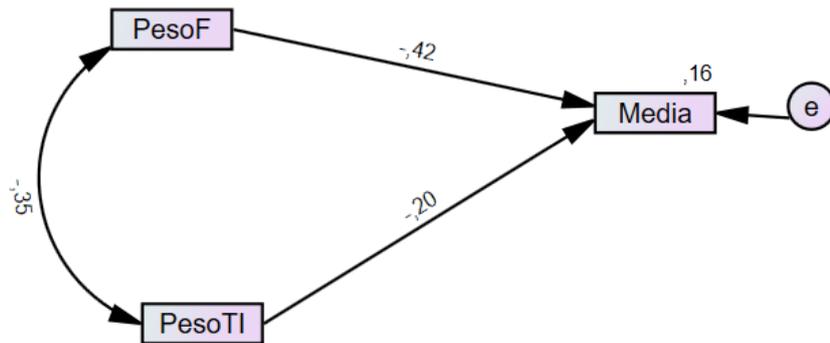


Figura 4.12: Modelo 2 Estudo 3

Entre todos os modelos apresentados este é o que tem uma variação explicada da variável média menor. Isto significa que as variáveis são apenas capazes de explicar o valor da média em 16%. Conclui-se ainda que o peso da frequência tem um impacto bastante negativo no valor da média uma vez que este apresenta um efeito direto negativo de - 42%. (Ver Figura 4.12)

Observando a Tabela 4.25 confirma-se que embora, para o PesoTI o valor do P não seja inferior a 0,001, pode ser considerado uma vez que é inferior a 0,005.

Tabela 4.25: Pesos Regressão Modelo 2 Estudo 3

			Estimativa	S.E.	C.R.	P
Media	<---	PesoF	-,015	,002	-6,628	***
Media	<---	PesoTI	-,009	,003	-3,108	,002

4.2.4 Discussão

Após o desenvolvimento da análise a partir de SEM, foi possível encontrar três modelos que têm uma variação de explicação da média de: (i) 27%, para a base de dados completa, isto é, para todos os valores da média; (ii) 25%, para a base de dados com valores de média iguais ou superiores a 10, e portanto, as médias com aprovação à respectiva UC; e, finalmente, (iii) 16% que corresponde à base de dados dos alunos com média igual ou superior a 14, porque o objetivo do estudo era encontrar um modelo onde os alunos tivessem o melhor proveito possível e uma vez que a partir de 14 valores já é um bom resultado optou-se por utilizar essa base de dados.

Infelizmente isto significa que os modelos não são capazes de explicar todo o sucesso escolar, apenas uma percentagem, e portanto, existem outros fatores externos que explicam os restantes 75% e 80%.

Segundo Afonso (2010), fatores como a escolaridade dos pais, o rendimento do agregado familiar, categoria socioprofissional de classe do agregado familiar, o género, a autorregulação, a autoeficácia da aprendizagem, a motivação para o estudo, o relacionamento com os colegas e o relacionamento com os professores são também fatores explicativos para o sucesso escolar. Afonso (2010) conclui que o modelo sociopsicológico com os fatores: classificação do relacionamento com os professores, motivação para o estudo, género feminino e rendimento do agregado familiar são os fatores mais importantes para o sucesso académico concluindo que essas variáveis têm uma variação explicada de 52%.

Concluindo os dois estudos juntos, o método de avaliação correto alinhado com todos os fatores sociopsicológicos, são capazes de explicar o sucesso académico em 77%.

4.3 Educational Data Mining

Neste subcapítulo serão abordadas as técnicas e os modelos de previsão com a linguagem R para prever como se comportaria a média dos alunos caso fossem utilizados os modelos obtidos no subcapítulo anterior (modelação de equações estruturais). A KDD foi a técnica escolhida para desenvolver esta previsão uma vez que esta tem como objetivo extrair conhecimento através dos dados. Este método difere dos restantes uma vez que não se foca em questões de negócio ou criação de modelos, mas sim na descoberta de conhecimentos a partir dos dados. Nesta dissertação foram testados dez modelos de previsão. Para que a análise seja mais completa e robusta, será também efetuada uma comparação entre o comportamento dos diferentes modelos caso fossem utilizadas todas as variáveis referentes ao método de avaliação.

Para que os modelos possam ser comparados e avaliados, serão utilizadas as seguintes métricas de avaliação: MAE, MAPE, RAE e NMAE, uma vez que estas são as métricas mais utilizadas para avaliar os modelos de previsão de problemas de regressão (Silva, Moro, Rita, & Cortez, 2018).

MAE, conhecida por Erro Médio Absoluto, é uma das métricas mais utilizadas para avaliar a precisão dos modelos de previsão. Esta métrica consiste na média da diferença absoluta entre o seu valor real ($Verd_i$) para todas as n observações e o total dos valores previstos ($Prev_i$) para uma dada variável de output. Avalia o desvio da capacidade de previsão do algoritmo. E é dada pela seguinte expressão (Silva, Moro, Rita, & Cortez, 2018):

$$MAE = \frac{1}{n} \sum_{i=1}^n |Verd_i - Prev_i|$$

MAPE, o Erro Percentual Absoluto Médio, é uma medida de precisão da previsão de um modelo de previsão em estatística. E é basicamente o rácio do *MAE* dividido pelo total dos valores reais ($Verd_i$). Esta métrica é dada pela seguinte expressão (Silva, Moro, Rita, & Cortez, 2018):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|Verd_i - Prev_i|}{Verd_i}$$

RAE ou Erro Relativo Absoluto é a diferença entre os valores totais verdadeiros e os valores totais previstos a dividir pela diferença entre os valores médios totais (Avg_i) e os valores totais previstos (Silva, Moro, Rita, & Cortez, 2018).

$$RAE = \frac{1}{n} \sum_{i=1}^n \frac{|Verd_i - Prev_i|}{|Avg_i - Prev_i|}$$

NMAE, Erro Médio Absoluto Normalizado, é a distribuição do *MAE* através da diferença entre os valores máximo (max) e mínimo (min) da variável output (Silva, Moro, Rita, & Cortez, 2018).

$$NMAE = \frac{MAE}{max - min}$$

Uma vez que todas as métricas utilizadas representam valores de erro, é importante que esse valor seja o mais baixo possível de forma a que o modelo tenha melhor desempenho na avaliação da previsão.

Importa ainda referir que o método de partição utilizado para a previsão foi o *k-fold*, com k igual a 10. Isto porque o método de partição *holdout* é utilizado para quando existe um grande volume de dados, o que não acontece nestes estudos.

4.3.1 Estudo 1

O Estudo 1 diz respeito ao estudo dos diversos modelos de previsão utilizados para prever a variável média, tendo em conta todas as variáveis “peso” presentes na base de dados inicial.

Na Tabela 4.26 observar-se que os modelos com melhores resultados nas métricas são os seguintes: randomForest e KSVM. No entanto, à exceção dos modelos RVM e XGBOOST, todos os modelos têm uma boa avaliação nas quatro métricas estudadas.

Tabela 4.26: Estudo 1 Métricas

Métricas/ Modelos	MAE	MAPE	NMAE	RAE
RANDOMFOREST	0,8033	6,6247	7,6413	62,0372
KSVM	0,8441	6,9949	8,0294	65,1879
MLPE	0,8537	7,0040	8,1207	65,9294
MLP	0,8821	7,2050	8,3910	68,1236
MARS	1,0295	8,3902	9,7925	79,5024
RPART	1,0295	8,3902	9,7925	79,5024
MR	1,0475	8,5287	9,9636	80,8917
CV.GLMNET	1,1110	9,0665	10,5678	85,7968
RVM	4,0764	31,2981	38,7759	314,8090
XGBOOST	6,1407	46,5537	58,4115	474,2245

É possível concluir-se então, que estes são os dois modelos que para este estudo apresentam melhores resultados de previsão.

No que concerne ao nível de importância, as Figuras seguintes (Figura 4.13 e 4.14) apresentam uma representação gráfica do nível de importância atribuído por cada modelo a cada variável “peso”. (Nas Figuras do Anexo C podem ser consultados os gráficos de importância dos restantes modelos).

O primeiro gráfico (Figura 4.13) apresentado diz respeito ao modelo RANDOMFOREST. Observa-se que este modelo atribui uma grande importância à variável peso da frequência (PesoF) com um valor de aproximadamente 21% em segundo lugar destaca-se o PesoTG com cerca de 18,5%. Observa-se ainda que o Peso2T, peso de 2 testes, é a variável com menor importância para este modelo com menos de 5% de importância.

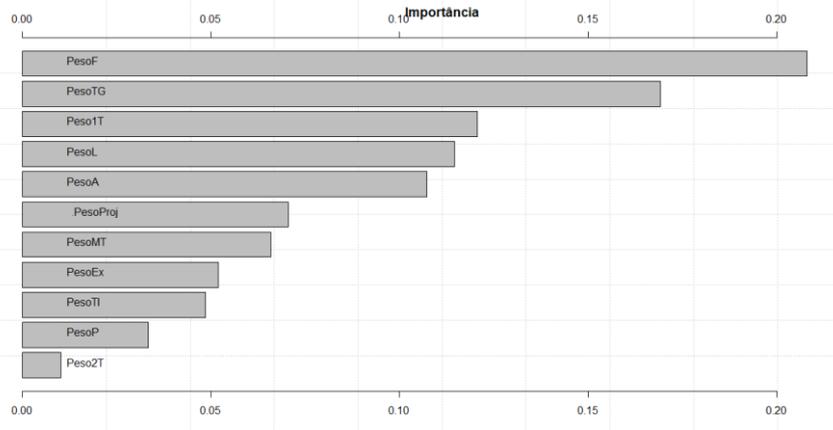


Figura 4.13: Gráfico de Importância Estudo 1 randomForest

O gráfico seguinte (Figura 4.14) representa os níveis de importância atribuídos a cada variável numérica. Pelo modelo KSVM verifica-se que este atribui maior importância às variáveis PesoTI e PesoEx com cerca de 11,5% e 10,5% respetivamente. No entanto observa-se que este modelo atribui níveis de importância aproximados às variáveis, isto é, a diferença entre a variável com mais importância e com menos importância difere apenas em aproximadamente 5%, não existindo uma diferença muito acentuada entre as variáveis.

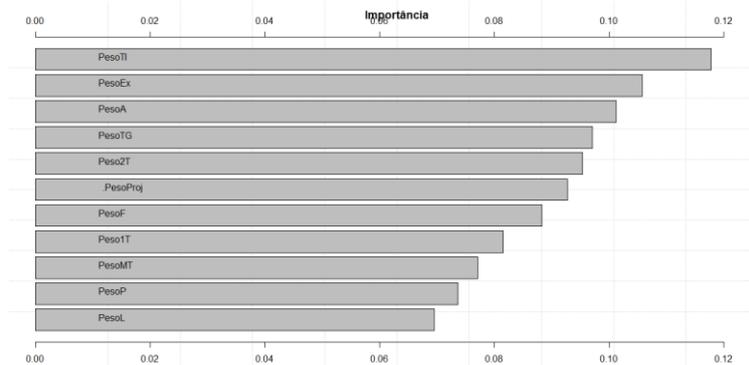


Figura 4.14: Gráfico de Importância Estudo 1 KSVM

Embora os modelos com melhores avaliações de previsão tenham sido os acima referidos, apresenta-se o gráfico da Figura 4.15 que contém a média calculada através dos outputs de todos os modelos estudados bem como uma comparação com a média real.

Observa-se que os dois modelos com piores avaliações são os modelos com a média das médias previstas mais baixas e bastante longe da média real. Também a média das médias prevista do modelo MARS é inferior à média das médias reais. Em geral, todos os modelos, à exceção dos referidos anteriormente têm valores de média prevista semelhantes à média real. Observa-se ainda que os modelos eleitos pelas métricas, o KSVM e RANDOMFOREST, apresentam 13,027 e

12,981 respetivamente. A partir desta análise conclui-se que o modelo que apresenta maior média prevista é o KSVM.

MÉDIAS DOS ALGORITMOS ESTUDO 1

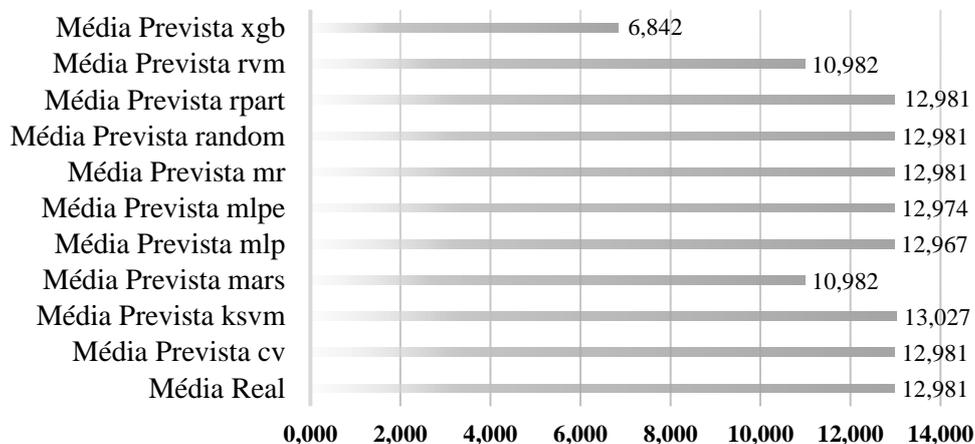


Figura 4.15: Gráfico de Médias Modelos Estudo 1

Para uma melhor compreensão do funcionamento dos modelos utilizados no estudo, apresenta-se uma breve análise ao gráfico de máximos e mínimos dos valores média (Figura 4.16). De acordo com a Figura 4.16, embora o modelo MARS tenha tido uma boa avaliação na previsão e a média das médias previstas tenha estado dentro dos parâmetros, este não é capaz de fazer uma previsão correta das médias uma vez que o valor máximo deste modelo é 28,496. Este valor não é possível na escala utilizada atualmente pela instituição, uma vez que os valores deveriam estar compreendidos entre 0 e 20 valores. Excluindo-se assim também este modelo como um bom modelo de previsão para este estudo. Observa-se ainda que nenhum dos modelos foi capaz de prever um valor máximo superior ao real, no entanto para todos os modelos aptos, os valores mínimos foram capazes de superar o valor mínimo real. Sendo o RANDOMFOREST e o CV. GLMNET os modelos com mínimos maiores.

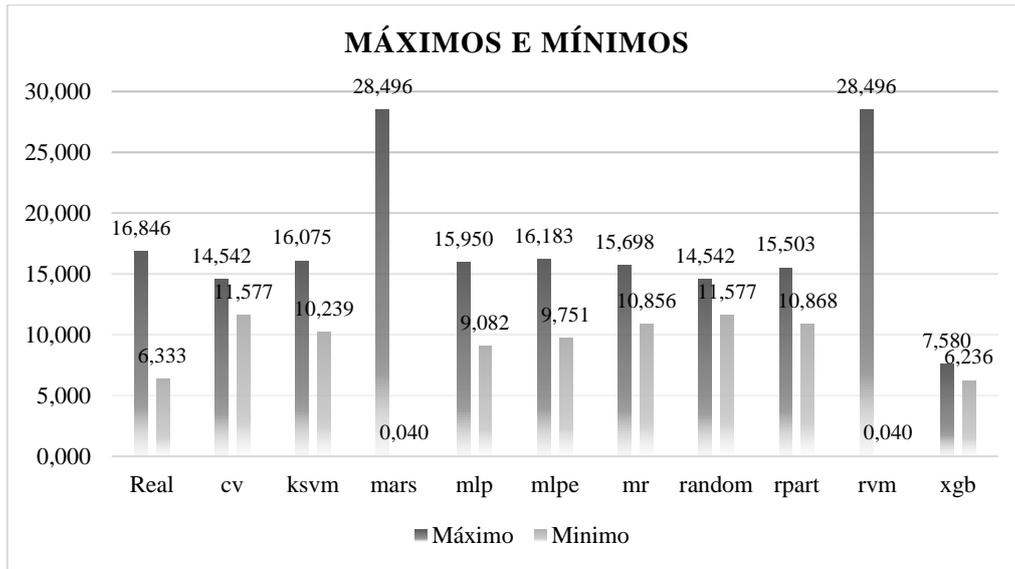


Figura 4.16: Máximos e Mínimos Estudo 1

4.3.2 Estudo 2

Após o desenvolvimento da análise apresentada em 4.2.1, onde foi possível perceber quais as variáveis “peso” que explicavam melhor a variável output média concluiu-se que as variáveis PesoL, PesoTG, PesoTI, PesoA e PesoProj são as variáveis com maior impacto na variável média. Desse modo, este segundo estudo representa uma avaliação dos 10 modelos testados onde a base de dados input é constituída pelas variáveis obtidas a partir das equações estruturais. Para esta análise não houve qualquer restrição à média contendo assim todo o tipo de valores possíveis incluindo valores onde a média se situa abaixo de 10.

Para este Estudo, segundo a Tabela 4.27, os modelos com melhor avaliação, e consequentemente com melhores previsões são os seguintes: KSVM e MLPE. As Figuras que se seguem (Figuras 4.17 e 4.18) representam os níveis de importância que cada modelo com melhor avaliação atribui às variáveis input. (Nas Figuras do Anexo D podem ser consultados os gráficos de importância dos restantes modelos).

Tabela 4.27: Estudo 2 Métricas

Métricas / Modelos	MAE	MAPE	NMAE	RAE
KSVM	0,93581	7,70926	8,90160	72,26932
MLPE	1,02333	8,35937	9,73413	79,02836
MLP	1,02355	8,36355	9,73621	79,04523
RPART	1,02771	8,38624	9,77581	79,36676
MARS	1,04820	8,55433	9,97066	80,94868
MR	1,06835	8,71099	10,16235	82,50489
RANDOMFOREST	1,08934	8,91825	10,36205	84,12619
CV.GLMNET	1,10452	9,02215	10,50642	85,29831
RVM	5,12158	39,96106	48,71742	395,52150
XGBOOST	6,13407	46,44878	58,34842	473,71250

Começando pelo modelo KSVM que atribui o maior nível de importância ao peso do trabalho de grupo, PesoTG, com cerca de 25% de importância, seguindo-se a variável peso do trabalho individual, PesoTI com aproximadamente 22%. Sendo a variável com menos importância a PesoL (peso dos laboratórios) com cerca de 14%, o que poderá ser um ponto positivo para este modelo uma vez que o PesoL apresenta uma relação inversa com a variável output média. (Ver Figura 4.17)

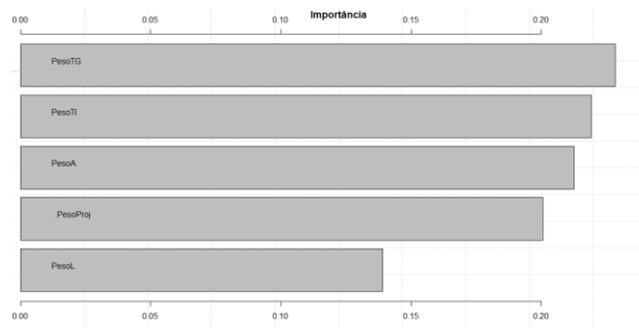


Figura 4.17: Gráfico de Importância Estudo 2 KSVM

Observando os níveis de importância do modelo MLPE na Figura 4.18, verifica-se que também neste modelo, o PesoTG é a variável com maior importância. Tal como o modelo anterior este também atribui o seu nível de menor importância à variável PesoL (peso dos laboratórios). Observa-se ainda que as variáveis PesoTG e PesoProj juntas têm mais de 50% de nível de importância para o MLPE.

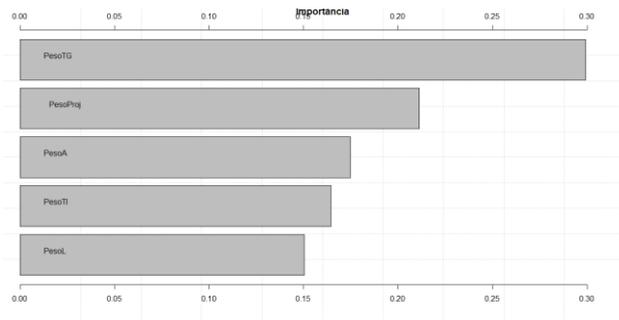


Figura 4.18: Gráfico de Importância Estudo 2 MLPE

No gráfico seguinte (Figura 4.19) verifica-se que em geral todos os modelos têm médias das médias previstas equivalentes à média das médias reais. Os modelos com pior avaliação de previsão são aqueles que contêm médias mais baixas e os modelos escolhidos por terem uma melhor avaliação de previsão, KSVM e MLPE, têm médias previstas de 13,013 e 12,981 valores respetivamente.

MÉDIAS DOS ALGORITMOS ESTUDO 2

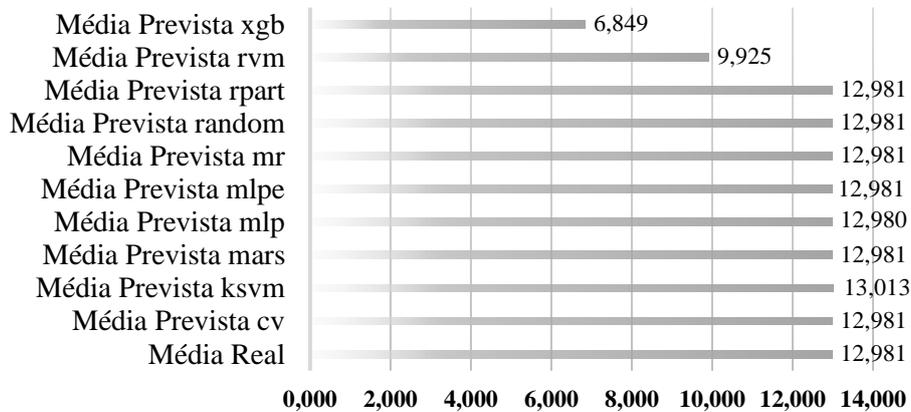


Figura 4.19: Gráfico de Médias Modelos Estudo 2

Comparando com o estudo anterior, conclui-se que as médias previstas pelos mesmos modelos mas com variáveis diferentes não diferem muito entre si.

Observando a Figura 4.20 que contém os valores máximos e mínimos do segundo estudo. Observa-se que à semelhança com o estudo anterior que nenhum dos modelos foi capaz de fazer previsões de média máximas superiores à média máxima real, em relação ao valor de médias mínimas em geral todos os modelos foram capazes de prever valores superiores.

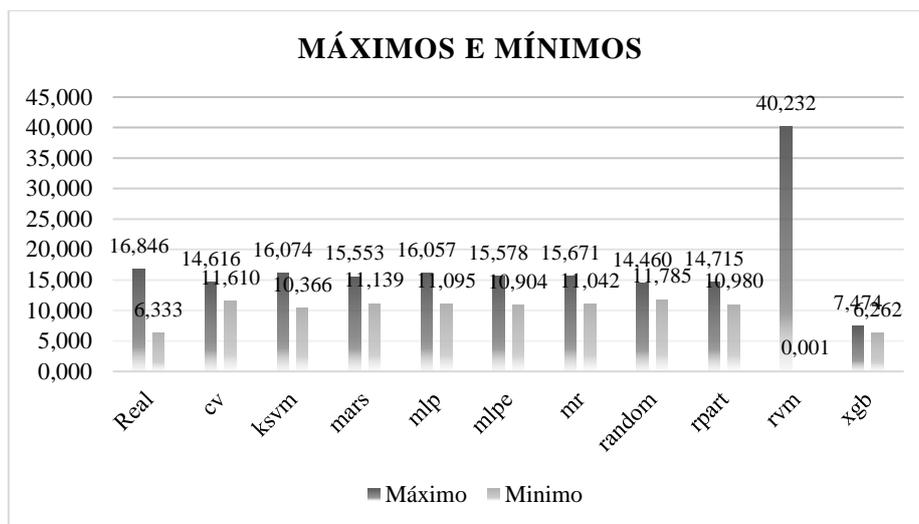


Figura 4.20: Máximos e Mínimos Estudo 2

4.3.3 Estudo 3

As figuras apresentadas abaixo (Figura 4.21 e Figura 4.22) são uma representação da importância das diversas variáveis para os modelos com melhor avaliação segundo as métricas de avaliação da previsão, MAE, MAPE, NMAE e RAE. Para estes testes, a base de dados input utilizada contém as seguintes variáveis: PesoTG, PesoTI, PesoL, PesoA e PesoL. Salienta-se que, para este terceiro estudo, foi importante considerar-se apenas as médias onde os alunos tiveram sucesso, isto é, foram consideradas apenas médias iguais ou superiores a 10.

Na Tabela 4.28 observa-se que os modelos com melhores resultados para este teste são o KSVM e o RPART, embora todos os modelos, à exceção dos modelos RVM e XGBOOST, apresentem bons valores para as métricas MAE, MAPE, NMAE e RAE. (Nas figuras do Anexo E podem ser consultados os gráficos de importância dos restantes modelos).

Tabela 4.28: Estudo 3 Métricas

Métricas/ Modelos	MAE	MAPE	NMAE	RAE
KSVM	0,8450	6,5011	12,4458	72,5273
RPART	0,9232	7,1482	13,5972	79,2368
MLPE	0,9317	7,2045	13,7233	79,9718
MLP	0,9323	7,2087	13,7313	80,0185
MARS	0,9448	7,2958	13,9160	81,0946
MR	0,9766	7,5427	14,3840	83,8221
RANDOMFOREST	0,9869	7,6414	14,5355	84,7052
CV.GLMNET	1,0059	7,7763	14,8155	86,3364
XGBOOST	6,2146	46,6366	91,5317	533,3967
RVM	8,1577	62,6304	120,1512	700,1751

Começando pelo modelo KSVM, Figura 4.21, uma vez que foi o modelo que apresentou os melhores resultados para as métricas de avaliação. Observa-se que o KSVM atribui maior importância ao peso do trabalho de grupo com um nível de importância superior a 20%. No entanto à exceção da variável PesoL, que apresenta um valor aproximado de 12% de importância, todas as variáveis apresentam valores de importância de entre os 20% e os 23%.

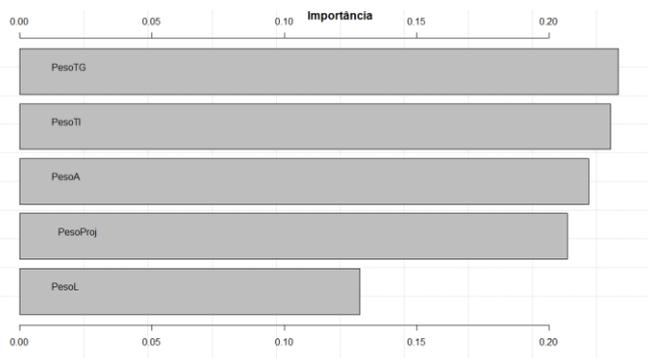


Figura 4.21: Gráfico de Importância Estudo 3 KSVM

A Figura 4.22 apresenta o gráfico de importância para o modelo Rpart, este atribui mais de 50% de importância a uma só variável sendo esta o peso do trabalho de grupo. As restantes variáveis têm todas níveis de importância inferiores a 15%. Sendo a variável trabalho individual com menor importância para o modelo.

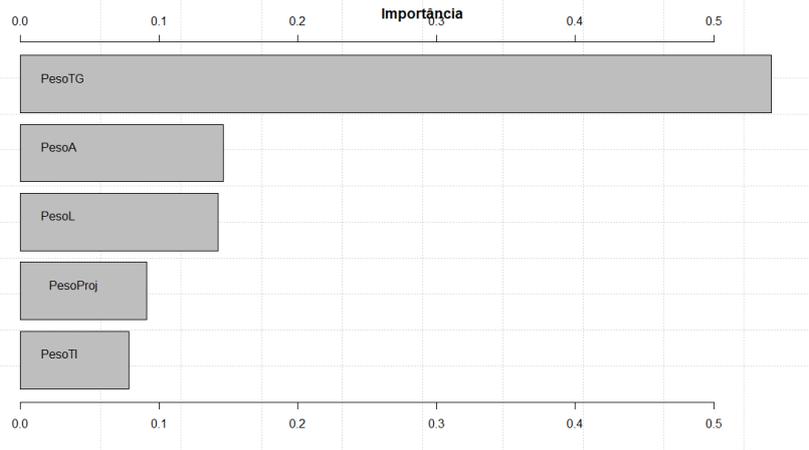


Figura 4.22: Gráfico de Importância Estudo 3 Rpart

No gráfico da Figura 4.23 é possível observar que as médias previstas em geral são semelhantes à média real das médias das diferentes UCs. Importa salientar que para este teste não foram contabilizadas médias negativas e por esse motivo as médias são superiores às observadas nos dois estudos anteriores. Os modelos XGBOOST e RVM são mais uma vez os que apresentam os piores valores de média prevista. As médias similares significam que os modelos são capazes de prever a média de forma aproximada da real considerando apenas as variáveis selecionadas. Os modelos com melhor avaliação de previsão para este estudo têm médias previstas de 13,096 e 13,159, sendo estas inferiores à média real.

MÉDIAS DOS ALGORITMOS ESTUDO 3

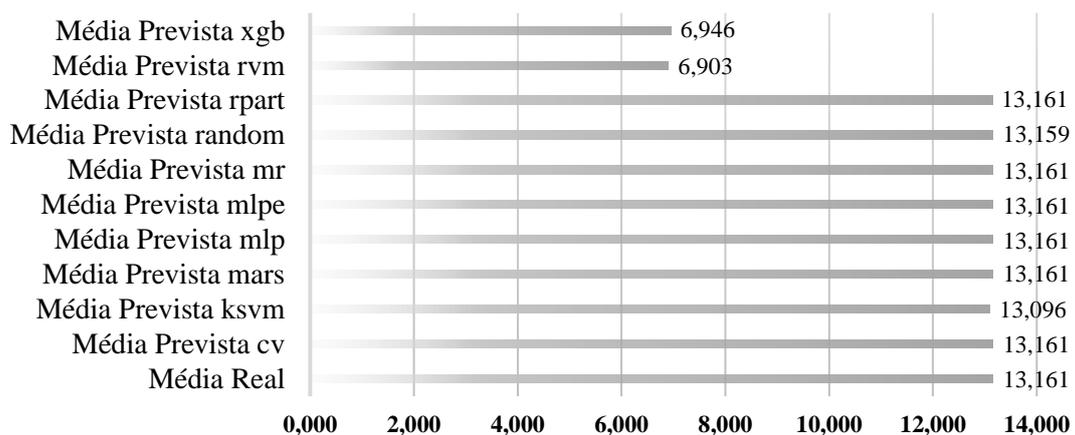


Figura 4.23: Gráfico de Médias Modelos Estudo 3

A Figura 4.24 apresenta o gráfico de médias máximas e médias mínimas referentes ao Estudo 3. Os modelos RVM e XGBOOST foram excluídos pelos mesmos motivos que nos estudos anteriores.

Relativamente à análise dos valores máximos de média, observa-se que em nenhum dos modelos o máximo previsto é capaz de superar o valor máximo das médias reais. No caso dos valores mínimos, à semelhança dos dois estudos anteriores, os valores previstos são superiores ao valor mínimo real para todos os modelos.

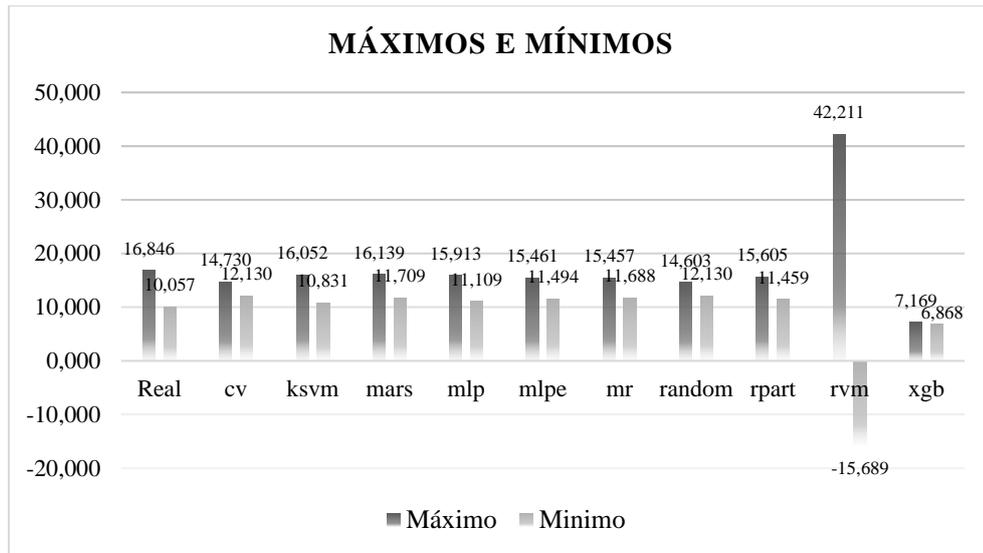


Figura 4.24: Máximos e Mínimos Estudo 3

4.3.4 Estudo 4

Os gráficos e tabela que se seguem representam uma análise aos mesmos 10 modelos utilizados nos estudos anteriores mas com a diferença das variáveis input e valores da média. Para estes testes foram consideradas apenas as variáveis PesoF e PesoTI porque segundo o modelo desenvolvido a partir das equações estruturas são estas as variáveis com maior percentagem de explicação da variável média para valores iguais ou superiores a 14.

Para este estudo e observando a Tabela 4.29 conclui-se que os modelos com melhores valores de MAE, MAPE, NMAE e RAE são os modelos KSVM e RANDOMFOREST. Observa-se que para todas as métricas estes são os modelos com valores mais baixos e consequentemente com a previsão do valor média mais preciso. (Nas Figuras do Anexo F podem ser consultados os gráficos de importância dos restantes modelos).

Tabela 4.29: Estudo 4 Métricas

Métricas/ Modelos	MAE	MAPE	NMAE	RAE
K SVM	0,3955	2,6255	13,8957	77,0529
RANDOMFOREST	0,3988	2,6653	14,0111	77,6933
REPART	0,4272	2,8570	15,0089	83,2260
MLP	0,4471	2,9931	15,7102	87,1148
MARS	0,4566	3,0507	16,0442	88,9665
MLPE	0,4586	3,0654	16,1134	89,3507
MR	0,4591	3,0676	16,1304	89,4447
CV.GLMNET	0,4881	3,2579	17,1512	95,1051
XGBOOST	7,0301	47,4183	247,0033	1369,6600
RVM	10,8391	73,0874	380,8321	2111,7550

Os gráficos seguintes (Figuras 4.25 e 4.26) representam o nível de importância para os modelos K SVM e RANDOMFOREST respetivamente. A partir destes gráficos verifica-se que os dois modelos atribuem maior importância ao peso da frequência (PesoF), com mais de 60%, e ao peso do trabalho individual (PesoTI) em relação às restantes variáveis.

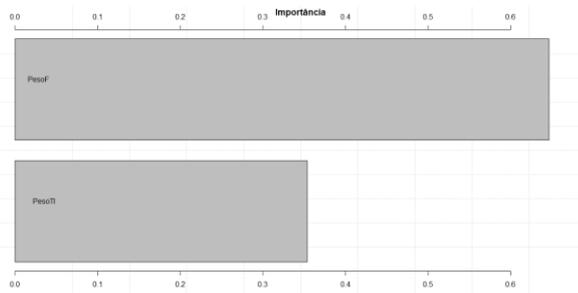


Figura 4.25: Gráfico de Importância Estudo 4 K SVM

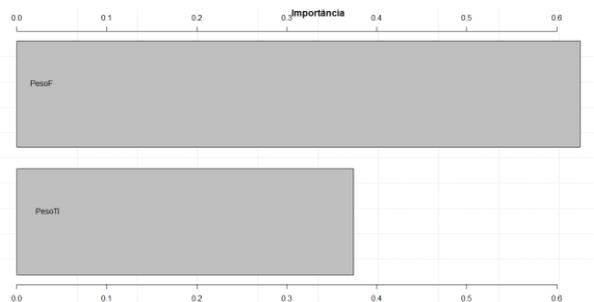


Figura 4.26: Gráfico de Importância Estudo 4 randomForest

A Figura 4.27 representa o gráfico das médias dos 10 modelos estudados. Novamente, à exceção dos modelos com pior avaliação de previsão, segundo as métricas todos os restantes previram médias semelhantes à média real, embora com apenas duas variáveis input estes modelos foram capazes de fazer uma previsão excelente. Entre todos os modelos, os seleccionados

com as melhores avaliações de previsão, KSVM e RANDOMFOREST, previram uma média de 14,720 e 14,799 respectivamente, sendo o segundo o modelo com maior média prevista. Não esquecendo que para este quarto estudo foram apenas consideradas as médias iguais ou superiores a 14.



Figura 4.27: Gráfico de Médias Modelos Estudo 4

O gráfico de máximos e mínimos referente ao quarto estudo é apresentado na Figura 4.28. Para este estudo as variáveis de input foram apenas as PesoF e PesoTI e apenas foram consideradas médias iguais ou superiores a 14 valores. Como é possível confirmar pelo gráfico mais uma vez e à semelhança com os estudos anteriores, todos os valores máximos previstos são inferiores ao valor máximo real. No entanto, os valores mínimos apresentados e previstos pelos modelos são na sua maioria superiores a 14 valores. Desta análise foram excluídos os modelos XGBOOST e RVM uma vez que estes não apresentam uma boa avaliação.

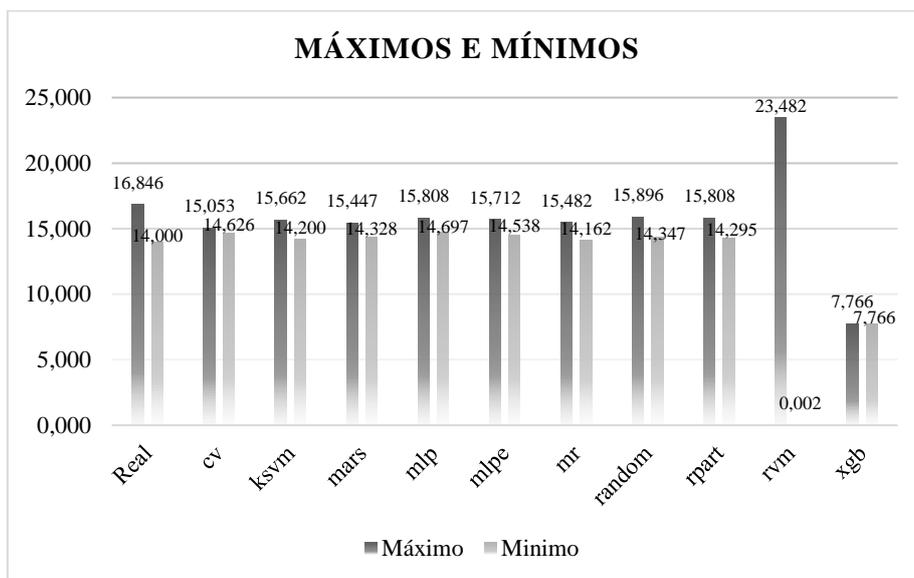


Figura 4.28: Máximos e Mínimos Estudo 4

4.3.5 Discussão

Neste capítulo verificou-se que, embora em todos os estudos um dos modelos com melhor avaliação tenha sido o KSVM, cada estudo contém avaliações de previsão diferentes para cada modelo. No entanto observando os gráficos das médias previstas, verifica-se que em geral todos os modelos independentemente das variáveis de input foram capazes de prever valores de média bastante aproximados à média real. Todos os modelos, à exceção do RVM e XGBOOST, têm uma boa avaliação de previsão dada pelas métricas e por isso são confiáveis.

O gráfico seguinte (Figura 4.29) representa uma comparação entre os Estudos 1 e 2. Uma vez que estes estudos apenas diferem nas variáveis input, era necessário compreender melhor como é que se baseava a previsão, bem como se havia alguma diferença devido ao input ser diferente. Uma vez que não é possível observar diretamente através dos gráficos média, apresenta-se o gráfico de máximos e mínimos (Figura 4.29).

A partir do gráfico ilustrado na Figura 4.29, é possível compreender que os valores não diferem muito entre modelos, no entanto existem pequenas diferenças. Começando pelos modelos com pior avaliação, observa-se que para os dois estudos o RVM atribui valores de média máxima impossíveis e o XGBOOST valores demasiado baixos. No entanto, no primeiro modelo referido a previsão do estudo 2 é bastante superior ao valor previsto do estudo 1 para valores de média máximos. Em relação ao modelo MARS no Estudo 1, este teve um valor máximo superior a 20, isto significa que não foi capaz de prever bem o valor média quando o input incluiu todas as variáveis recolhidas.

Em relação aos restantes modelos, observa-se que em geral todos tiveram valores de média máxima superiores no Estudo 2, mesmo que só difiram em décimas. O mesmo acontece para os valores de média mínimos. Para valores de média mínimos destacam-se os modelos MLP, MLPE e MR que foram capazes de prever valores mais altos de mínimo no Estudo 2 do que no Estudo 1 (sendo um valor a mais de diferença). Isto pode significar que realmente utilizando apenas as variáveis selecionadas pelas equações estruturais como método de avaliação a previsão das notas mínimas dos alunos no ensino superior são mais elevadas. O mesmo não acontece para as notas máximas, diferindo apenas em algumas décimas.

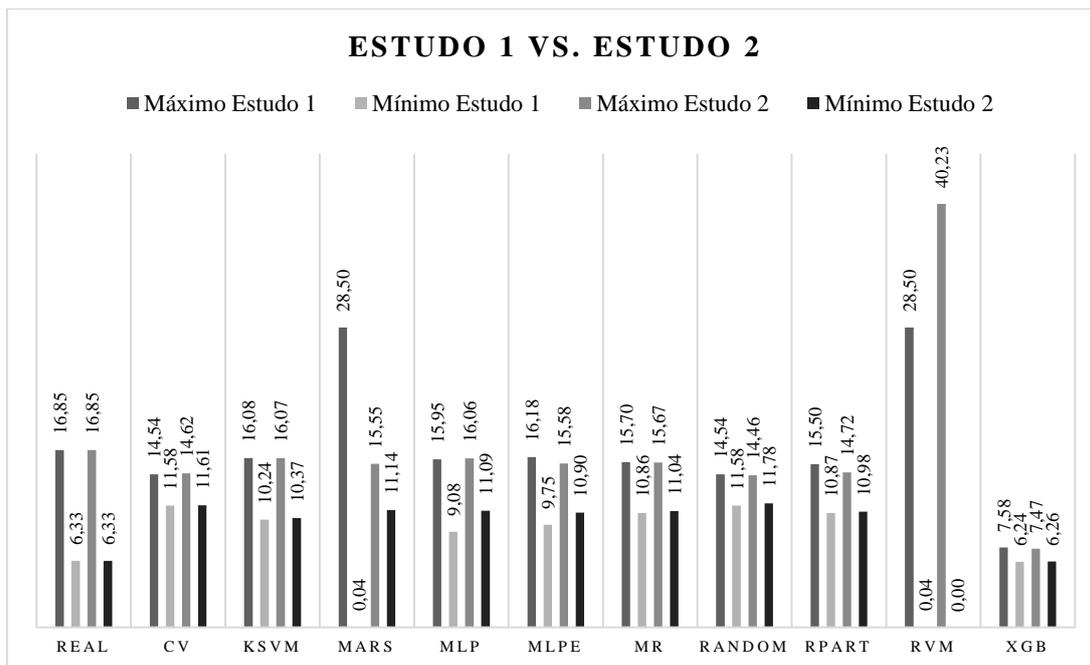


Figura 4.29: Máximos e Mínimos Estudo 1 vs. Estudo 2

Após a análise dos gráficos de máximos e mínimos de todos os estudos verifica-se também que, embora os valores máximos não sejam superiores ao valor máximo real, os valores mínimos em todos os casos são, o que mais uma vez vem sustentar a ideia inicial que embora as médias mais altas não possam ser afetadas pela implementação dos novos métodos de avaliação as médias mais baixas podem.

No entanto no caso de médias superiores a 14, em que as variáveis com maior explicação são o peso da frequência e o trabalho individual é necessário ter em conta que as médias mais altas, superiores a 16, podem ser afetadas se apenas forem considerados estes dois métodos de avaliação.

4.4 Estudo Complementar

Os dados obtidos permitiram ainda uma análise complementar relativamente à área de estudos das UCs e ao ano letivo. As UCs foram agrupadas em três grandes grupos: “Ciências empresariais e Direito”, “Matemática” e “Tecnologias”.

De acordo com a Tabela 4.30, que indica quantos casos existem em cada grupo, o grupo que contém mais registos é o de “Ciências empresariais e Direito”, tendo cerca de 57,8% dos registos.

Tabela 4.30: Resumo do processamento de casos

Resumo do processamento de casos							
codigos		n	Percentagem	n	Percentagem	n	Percentagem
Media	Ciências empresariais e direito	458	100,0%	0	0,0%	458	100,0%
	Matemática	92	100,0%	0	0,0%	92	100,0%
	Tecnologias	242	100,0%	0	0,0%	242	100,0%

A hipótese de investigação colocada é a existência de diferenças entre as áreas científicas relativamente às médias dos alunos.

Sendo as hipóteses estatísticas as seguintes:

H_0 : Não existem diferenças entre os grupos “código da cadeira” relativamente à média.

H_a : Existem diferenças entre os grupos “código da cadeira” relativamente à média.

Para começar a análise é importante saber se as variáveis seguem a normalidade para saber que tipo de teste deve ser aplicado para o estudo.

Importa ainda referir que o estudo é feito para 3 amostras independentes com um $n \geq 30$.

Uma vez que a dimensão de cada amostra independente é superior a 30 observações recorrer-se-á ao teste do K-S (Kolmogorov-Smirnova) para analisar a normalidade. No grupo “Ciências empresariais e Direito” e “Matemática” o valor-p é superior a 0,005 o que significa que à partida estes dois grupos não seguem a distribuição normal. No grupo “Tecnologias” verifica-se que os dados seguem a distribuição normal porque $\text{valor-p} = 0,001 < 0,005$ (Ver Tabela 4.31).

No entanto, para aplicar o método da ANOVA, todas as variáveis deveriam seguir a distribuição normal.

Tabela 4.31: Teste à Normalidade

codigos		Teste da Normalidade					
		Kolmogorov-Smirnova			Shapiro-Wilk		
Media		Estatísticas	df	Sig.	Estatísticas	df	Sig.
	Ciências empresariais e direito	0,033	458	,200*	0,993	458	0,028
	Matemática	0,098	92	0,030	0,984	92	0,335
	Tecnologias	0,078	242	0,001	0,977	242	0,001

Passando agora ao segundo teste à normalidade em que é necessário observar os valores da Assimetria e Curtose dos grupos “Ciências empresariais e Direito” e “Matemática”. Não é necessário observar os mesmos valores para o grupo “Tecnologias” uma vez que já se verificou que este segue a distribuição normal. Calculando como anteriormente no subcapítulo análise de dados verifica-se igualmente que nenhum dos grupos (“Ciências empresariais e Direito” e “Matemática”) seguem a distribuição normal uma vez que o rácio entre SkewnessStatistic e SkewnessStd. Erro e o rácio entre KurtosisStatistic e KurtosisStd. Erro não estão compreendidos entre -2 e 2. (Ver Tabela 4.32)

Tabela 4.32: Assimetria e Curtose

		Descritivas		
códigos			Estatísticas	Erro Padrão
Media	Ciências empresariais e direito	Média	13,254	0,065
		Skewness	-0,268	0,114
		Kurtosis	0,192	0,228
	Matemática	Média	12,263	0,119
		Skewness	-0,391	0,251
		Kurtosis	0,774	0,498
	Tecnologias	Média	12,737	0,131
		Skewness	-0,494	0,156
		Kurtosis	-0,075	0,312

Observa-se ainda que existe uma diferença entre os grupos que pode ser observada nas Figuras 4.30 e 4.31 abaixo.

Na Figura 4.30 verifica-se que o grupo de UCs com os piores resultados pertence ao grupo de matemática e o melhor pertence ao grupo de ciências empresariais e direito.

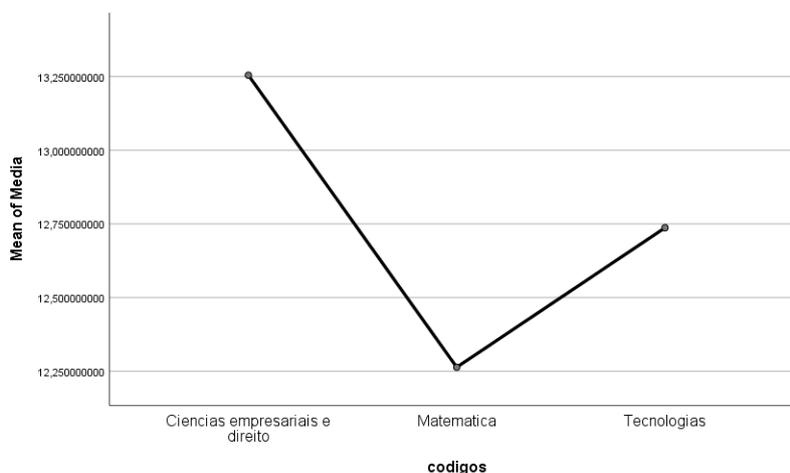


Figura 4.30: Gráfico Médias por grupo

Pelos motivos apresentados acima, o teste utilizado para estudar a hipótese de investigação é o de Kruskal-Wallis.

Ao observar a Tabela 4.33 confirma-se que o valor-p é inferior a 0,001 e consequentemente inferior a 0,005 significa que existem evidências estatísticas suficientes para afirmar a existência de diferenças estatísticas significativas por área científica relativamente à média, assumindo um nível de significância de 5%.

Tabela 4.33: Teste Estatístico Kruskal-Wallis

Teste Estatístico		
	Media	Media (+10)
Kruskal-Wallis H	37,810	22,902
df	2	2
Asymp. Sig.	<0,001	<0,001

Na figura 4.31 comprova-se a existência das diferenças entre as áreas científicas.

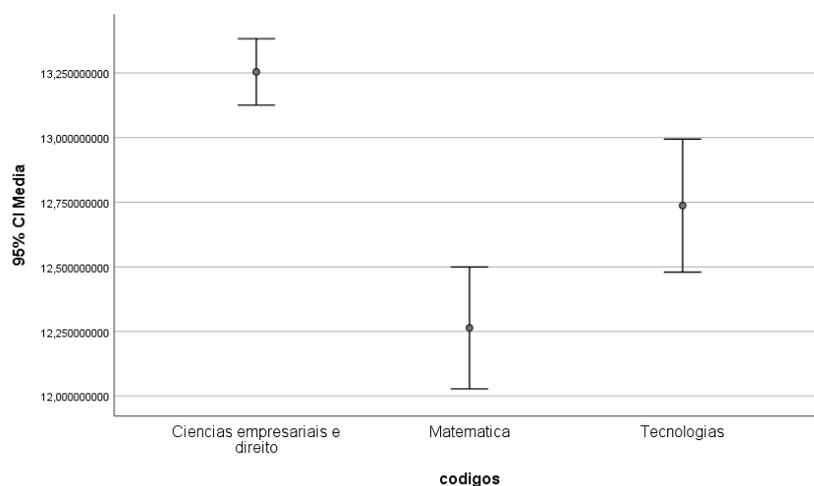


Figura 4.31:Boxplot areas científicas

A Figura 4.32 ilustra a evolução das médias ao longo do tempo. Observa-se que o primeiro ano do estudo é o que contém os valores de média mais baixos e o último ano o que contém os mais altos. Nota-se ainda que no ano letivo 2016/2017 existe uma ligeira quebra no valor das médias. É também possível observar que em nenhum dos anos letivos as médias foram inferiores a 10, isto significa que em termos gerais a maioria dos alunos teve um rendimento académico positivo.

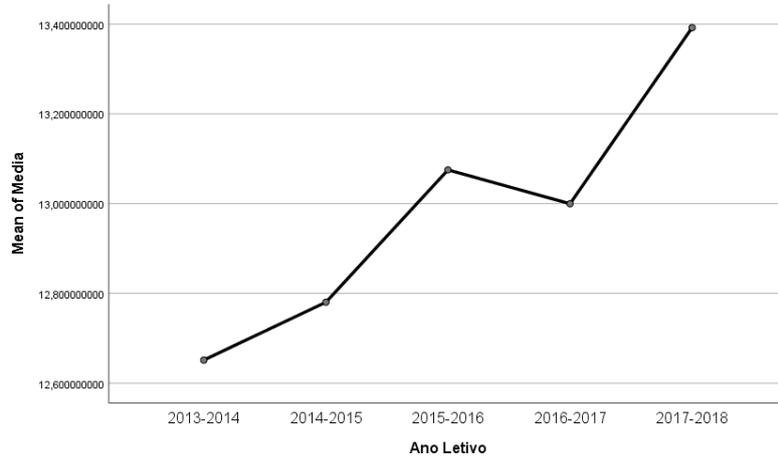


Figura 4.32: Gráfico Médias por ano letivo

5. Conclusão

Esta dissertação teve como objetivo principal, analisar a relação entre os métodos e instrumentos de avaliação e a classificação final das UCs. Foi possível concluir, que não existe uma forte relação mas que existe uma percentagem explicativa associada que difere nos casos em que se analisam (i) todas as médias, (ii) apenas as médias positivas e (iii) apenas as médias elevadas.

Foi possível verificar que para cada caso (i, ii e iii) existe um modelo estrutural diferente e que cada modelo tem uma percentagem distinta de explicação da variável média. Após este estudo ser realizado, foi ainda possível verificar, como se comportariam as médias dos alunos caso fossem aplicados exclusivamente os métodos de avaliação retirados a partir da modelação de equações estruturais.

Tal como Pimentel e Omar (2006), o objetivo desta dissertação era tentar perceber, como é que os métodos de avaliação podem influenciar os resultados dos alunos através de Data Mining. No entanto este trabalho foi além do estudo dos autores uma vez que foi verificado, recorrendo a equações estruturais, o impacto que os métodos de avaliação podem ter na média dos alunos, bem como a previsão tendo em conta os diferentes grupos (i, ii e iii).

Em suma, e respondendo a cada questão de investigação:

- Será que existe relação entre as variáveis “método de avaliação” e os resultados dos alunos?

Sim, concluiu-se que existem fortes indícios de relação embora essa relação não tenha muito peso.

- Em que medida estas variáveis são capazes de explicar a variável média?

A resposta dependerá do tipo de médias que estão a ser trabalhadas. Concluiu-se que quando se trata do total de médias da amostra (positivas e negativas), o valor de explicação da variável consegue alcançar os 27%. No entanto, no caso de médias apenas positivas o valor explicado é apenas de 25% e este valor de explicação reduz drasticamente quando se trata de médias consideradas boas, superiores a 14, com uma taxa de explicação apenas de 16%.

Alcançado este conhecimento foi possível, aplicando mecanismos de Data Mining, perceber-se como é que as médias poderiam ser afetadas. Concluiu-se que para as médias mais altas os modelos de previsão não foram capazes de prever valores superiores aos reais, no entanto quando se trata de médias mais baixas, observou-se que estes modelos foram capazes de prever, em alguns casos, valores superiores ao valor mínimo das médias reais. Por este motivo, é esperado que nas

UCs com médias mais baixas possam elevar as suas médias com os modelos de avaliação aqui propostos.

Em relação à última questão de investigação:

- Será que existem diferenças entre as áreas das Unidades Curriculares (Código da Cadeira) relativamente às médias das Unidades Curriculares?

Concluiu-se através de testes estatísticos que sim.

Como limitação durante o desenvolvimento desta dissertação observou-se que um dos problemas detetado quando a recolha dos dados foi iniciada foi a impossibilidade de se saber, se as notas presentes na ferramenta de Miguel et al. (2020) pertenciam a uma época de avaliação específica (primeira, segunda ou época especial) ou ao conjunto de todas as épocas. No entanto para este estudo era essencial que as notas fossem de primeira época, uma vez que o objetivo era estudar que método de avaliação teria mais impacto e seria melhor para a média dos alunos, uma vez que, habitualmente a segunda época e a época especial são avaliadas através de um exame. Por este motivo, foram consideradas que todas as notas se tratavam de época normal e foi a partir desse pressuposto que foram calculadas as médias de cada UC.

Como trabalho futuro sugere-se a extensão desta dissertação a partir da análise de mais fatores exteriores que podem influenciar a média dos alunos no ensino superior.

Referências

- Afonso, C. M. (2010). *Carlos Miguel Afonso*.
- Baepler, P., & James Murdoch, C. (2010). Academic Analytics and Data Mining in Higher Education. *International Journal for the Scholarship of Teaching and Learning*, 4(2). <https://doi.org/10.20429/ijstl.2010.040217>
- Baker, R., Isotani, S., & Carvalho, A. (2011). Mineração de Dados Educacionais: Oportunidades para o Brasil. *Revista Brasileira de Informática Na Educação*, 19(02). <https://doi.org/10.5753/rbie.2011.19.02.03>
- Barbosa, J. (2008). *A Avaliação da Aprendizagem como Processo Interativo : Um Desafio para o Educador*.
- Black, P., & Wiliam, D. (1998). Inside the black box. Raising standards through classroom assessment. *Londres: King's College London School of Education*, 19(1), 136–137. <https://doi.org/10.1111/j.1949-3606.2010.00017.x>
- Bresfelean, V. P., Bresfelean, M., Ghisoiu, N., & Comes, C. A. (2008). Determining students' academic failure profile founded on data mining methods. *Proceedings of the International Conference on Information Technology Interfaces, ITI*, 317–322. <https://doi.org/10.1109/ITI.2008.4588429>
- Bryan, C., & Clegg, K. (2006). Introduction. In C. Bryan & K. Clegg (Orgs.). In *Innovative Assessment in Higher Education*.
- Castro, F., Vellido, A., Nebot, À., & Mugica, F. (2007). Applying data mining techniques to e-learning problems. *Studies in Computational Intelligence*, 62, 183–221. https://doi.org/10.1007/978-3-540-71974-8_8
- Chism, N. V. N., Angelo, T. A., & Cross, K. P. (1995). Classroom Assessment Techniques: A Handbook for College Teachers. *The Journal of Higher Education*, 66(1), 108. <https://doi.org/10.2307/2943957>
- Denise Alves Ferreira Amorim, L., Leovigildo Fiaccone, R., Antônio de Souza Teles Santos, C., Terezinha Lana Pimenta de Moraes, L., Fernandes de Oliveira, N., Barbosa Oliveira, S., & Nadya Lima dos Santos, T. (2012). *Modelagem com Equações Estruturais: Princípios Básicos e Aplicações*. Retrieved from <http://repositorio.ufba.br/ri/handle/ri/17684>
- Dutt, A., Ismail, M. A., & Herawan, T. (2017). A Systematic Review on Educational Data Mining. *IEEE Access*, Vol. 5, pp. 15991–16005. <https://doi.org/10.1109/ACCESS.2017.2654247>
- Fernandes, D., & Fialho, N. (2012). Dez anos de práticas de avaliação das aprendizagens no ensino superior: Uma síntese da literatura (2000-2009). In *Ensino superior: Inovação e qualidade na docência* (pp. 3693–3707).
- Fernandes, D., Rodrigues, P., & Nunes, C. (2012). *Uma Investigação em Ensino, Avaliação e Aprendizagens no Ensino Superior*. Retrieved from <http://repositorio.ul.pt/handle/10451/8999>
- Garcia, J. (2009). Avaliação e aprendizagem na educação superior. *Estudos Em Avaliação Educacional*, 20(43), 201. <https://doi.org/10.18222/ae204320092045>
- Gaulden. (2010). *Classroom Assessment Techniques taken*.
- Gijbels, D., Segers, M., & Struyf, E. (2008). Constructivist learning environments and the (im)possibility to change students' perceptions of assessment demands and approaches to

- learning. *Instructional Science*, 36(5-6 EFFECTS OF CONSTR), 431–443.
<https://doi.org/10.1007/s11251-008-9064-7>
- Gil, P. D., da Cruz Martins, S., Moro, S., & Costa, J. M. (2020). A data-driven approach to predict first-year students' academic success in higher education institutions. *Education and Information Technologies*, 1–26. <https://doi.org/10.1007/s10639-020-10346-6>
- Leal, L. C. (1997). *Exames: uma via a prosseguir?*
- Leitão, I. (2013). *Os Diferentes Tipos de Avaliação : Avaliação Formativa e Avaliação Sumativa*. 10(1), 290–313. Retrieved from https://dre.pt/web/guest/pesquisa/-/search/115652962/details/normal?q=Decreto-Lei+n.º55%2F2018%0Ahttps://www.researchgate.net/publication/339956031_Avaliacao_Sumativa
- Maffezzolli, A. P., & Morais Pereira, T. de. (2020). AVALIAÇÃO DA APRENDIZAGEM NO ENSINO SUPERIOR: CONCEPÇÕES E DESAFIOS. *REVISTA APROXIMAÇÃO — VOLUME 02. NÚMERO 03. — ABR/MAI/JUN 2020*, 8–13.
- Marôco, J. (2014a). *Análise de Equações Estruturais Fundamentos teóricos, Software e Aplicações*.
- Marôco, J. (2014b). *Análise Estatística com o SPSS Statistics (6ª ed.)*.
- Miguel, H. (2019). *Provision of Academic Data for Research: A Step for Academic Success (ISCTE)*. Retrieved from <http://hdl.handle.net/10071/20124>
- Miguel, H. G., Ramos, P., da Cruz Martins, S., & Costa, J. M. (2020). Provision of academic data for research: A step for academic success. *Education for Information, Preprint(Preprint)*, 1–19. <https://doi.org/10.3233/efi-190349>
- Miranda, M. (1982). Docimologia em perspectiva. 8(1), 39–69.
- Peña-Ayala, A. (2014). *Educational data mining: A survey and a data mining-based analysis of recent works* Keywords: Data mining Educational data mining Data mining profile Educational data mining approach pattern Pattern for descriptive and predictive educational data mining approaches. <https://doi.org/10.1016/j.eswa.2013.08.042>
- Pereira, D. R., & Flores, M. A. (2012). Percepções dos estudantes universitários sobre a avaliação das aprendizagens: um estudo exploratório. *Avaliação: Revista Da Avaliação Da Educação Superior (Campinas)*, 17(2), 529–556. <https://doi.org/10.1590/s1414-40772012000200012>
- Pestana, M., & Gageiro, J. (2014). *Análise de dados para ciências sociais 6ª edição. Edições*.
- Pimentel, E. P., & Omar, N. (2006). Descobrendo Conhecimentos em Dados de Avaliação da Aprendizagem com Técnicas de Mineração de Dados. In *Anais do Workshop de Informática na Escola (Vol. 1)*. <https://doi.org/10.5753/CBIE.WIE.2006.%P>
- Rasmden, P. (2003). *Learning to teach in higher education*.
- Romero, Cristobal, & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12–27. <https://doi.org/10.1002/widm.1075>
- Romero, Cristóbal, Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers and Education*, 51(1), 368–384. <https://doi.org/10.1016/j.compedu.2007.05.016>
- Rosado, A., & Silva, C. (2010). *CONCEITOS BÁSICOS SOBRE AVALIAÇÃO DAS APRENDIZAGENS*. Retrieved from

<http://home.fmh.utl.pt/~arosado/ESTAGIO/conceitos.htm>

Silva, A. T., Moro, S., Rita, P., & Cortez, P. (2018). Unveiling the features of successful eBay smartphone sellers. *Journal of Retailing and Consumer Services*, 43, 311–324.

<https://doi.org/10.1016/j.jretconser.2018.05.001>

Trotter, E. (2006). Student perceptions of continuous summative assessment. *Assessment and Evaluation in Higher Education*, 31(5), 505–521.

<https://doi.org/10.1080/02602930600679506>

Vandamme, J. -P., Meskens, N., & Superby, J. -F. (2007). Predicting Academic Performance by Data Mining Methods. *Education Economics*, 15(4), 405–419.

<https://doi.org/10.1080/09645290701409939>

Zaïane, O. R. (2001). *Web Usage Mining for a Better Web-Based Learning Environment*.

<https://doi.org/10.7939/R3736M20P>

[Esta página foi intencionalmente deixada em branco].

ANEXOS

Anexo A – Tabela Coeficientes Stepwise Estudo 1 (Modelação de EE)

		Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	12,441	0,074		168,160	0,000		
	PesoTG	0,032	0,003	0,357	10,724	0,000	1,000	1,000
2	(Constant)	12,277	0,075		163,610	0,000		
	PesoTG	0,036	0,003	0,406	12,351	0,000	0,957	1,045
	PesoProj	0,032	0,004	0,241	7,328	0,000	0,957	1,045
3	(Constant)	12,219	0,074		164,854	0,000		
	PesoTG	0,034	0,003	0,382	11,760	0,000	0,942	1,062
	PesoProj	0,034	0,004	0,249	7,716	0,000	0,956	1,046
	PesoA	0,042	0,007	0,188	5,916	0,000	0,979	1,021
4	(Constant)	12,383	0,078		158,071	0,000		
	PesoTG	0,030	0,003	0,334	10,123	0,000	0,879	1,138
	PesoProj	0,032	0,004	0,240	7,568	0,000	0,953	1,049
	PesoA	0,040	0,007	0,179	5,725	0,000	0,977	1,024
	PesoL	-0,034	0,006	-0,181	-5,629	0,000	0,926	1,080
5	(Constant)	12,288	0,080		154,273	0,000		
	PesoTG	0,031	0,003	0,344	10,556	0,000	0,875	1,142
	PesoProj	0,034	0,004	0,251	8,021	0,000	0,948	1,055
	PesoA	0,039	0,007	0,177	5,722	0,000	0,976	1,024
	PesoL	-0,031	0,006	-0,166	-5,230	0,000	0,918	1,089
	PesoTI	0,022	0,005	0,150	4,870	0,000	0,985	1,015
6	(Constant)	12,238	0,082		148,797	0,000		
	PesoTG	0,030	0,003	0,335	10,229	0,000	0,863	1,159
	PesoProj	0,034	0,004	0,253	8,101	0,000	0,947	1,056
	PesoA	0,039	0,007	0,173	5,623	0,000	0,974	1,026
	PesoL	-0,030	0,006	-0,161	-5,068	0,000	0,914	1,094
	PesoTI	0,022	0,005	0,151	4,922	0,000	0,985	1,015
	PesoP	0,019	0,008	0,073	2,376	0,018	0,968	1,033
7	(Constant)	12,182	0,086		141,637	0,000		
	PesoTG	0,031	0,003	0,344	10,451	0,000	0,847	1,181
	PesoProj	0,035	0,004	0,260	8,307	0,000	0,936	1,068
	PesoA	0,039	0,007	0,174	5,653	0,000	0,974	1,026

	PesoL	-0,033	0,006	-0,175	-5,411	0,000	0,877	1,140
	PesoTI	0,023	0,005	0,155	5,049	0,000	0,982	1,019
	PesoP	0,021	0,008	0,080	2,577	0,010	0,959	1,043
	PesoMT	0,011	0,005	0,069	2,162	0,031	0,899	1,112
8	(Constant)	12,150	0,087		139,012	0,000		
	PesoTG	0,031	0,003	0,352	10,634	0,000	0,834	1,198
	PesoProj	0,036	0,004	0,264	8,427	0,000	0,933	1,072
	PesoA	0,039	0,007	0,175	5,710	0,000	0,974	1,027
	PesoL	-0,032	0,006	-0,171	-5,275	0,000	0,873	1,146
	PesoTI	0,023	0,004	0,158	5,149	0,000	0,979	1,021
	PesoP	0,021	0,008	0,083	2,670	0,008	0,957	1,045
	PesoMT	0,011	0,005	0,071	2,216	0,027	0,899	1,113
	Peso2T	0,022	0,011	0,061	1,980	0,048	0,977	1,023

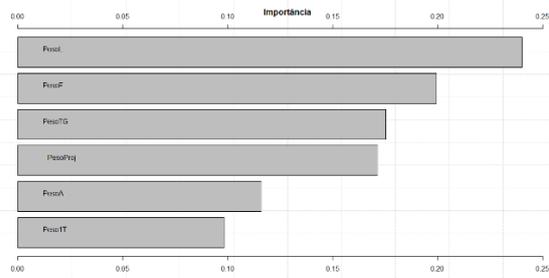
Anexo B – Tabela Coeficientes Stepwise Estudo 2 (Modelação de EE)

		Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	12,70	0,07		190,70	0,00		
	PesoTG	0,03	0,00	0,34	10,06	0,00	1,00	1,00
2	(Constant)	12,55	0,07		185,12	0,00		
	PesoTG	0,03	0,00	0,40	11,71	0,00	0,95	1,05
	PesoProj	0,03	0,00	0,24	7,17	0,00	0,95	1,05
3	(Constant)	12,49	0,07		186,72	0,00		
	PesoTG	0,03	0,00	0,37	11,17	0,00	0,94	1,06
	PesoProj	0,03	0,00	0,25	7,59	0,00	0,95	1,05
	PesoA	0,04	0,01	0,20	5,98	0,00	0,98	1,02
4	(Constant)	12,41	0,07		183,33	0,00		
	PesoTG	0,03	0,00	0,38	11,60	0,00	0,94	1,07
	PesoProj	0,03	0,00	0,26	8,08	0,00	0,95	1,06
	PesoA	0,04	0,01	0,19	5,98	0,00	0,98	1,02
	PesoTI	0,02	0,00	0,16	5,12	0,00	0,99	1,01
5	(Constant)	12,50	0,07		174,46	0,00		
	PesoTG	0,03	0,00	0,35	10,49	0,00	0,88	1,13
	PesoProj	0,03	0,00	0,26	7,99	0,00	0,94	1,06
	PesoA	0,04	0,01	0,19	5,85	0,00	0,98	1,02
	PesoTI	0,02	0,00	0,15	4,84	0,00	0,99	1,01
	PesoL	-0,02	0,01	-0,12	-3,52	0,00	0,93	1,07
6	(Constant)	12,43	0,08		162,25	0,00		
	PesoTG	0,03	0,00	0,37	10,86	0,00	0,84	1,19
	PesoProj	0,03	0,00	0,26	8,09	0,00	0,94	1,06
	PesoA	0,04	0,01	0,19	5,98	0,00	0,98	1,02
	PesoTI	0,02	0,00	0,16	5,09	0,00	0,98	1,02
	PesoL	-0,02	0,01	-0,10	-3,15	0,00	0,91	1,09
	PesoEx	0,02	0,01	0,09	2,66	0,01	0,94	1,07
7	(Constant)	12,38	0,08		157,03	0,00		
	PesoTG	0,03	0,00	0,36	10,58	0,00	0,83	1,21
	PesoProj	0,03	0,00	0,26	8,17	0,00	0,94	1,06

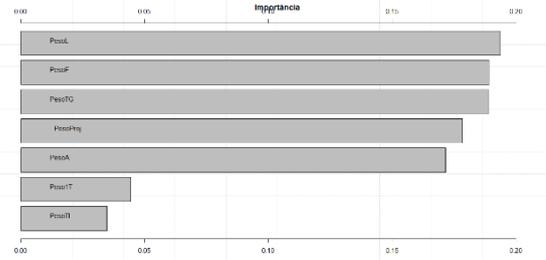
A influência dos métodos de avaliação nos resultados dos alunos do ensino superior

PesoA	0,04	0,01	0,19	5,89	0,00	0,97	1,03
PesoTI	0,02	0,00	0,16	5,14	0,00	0,98	1,02
PesoL	-0,02	0,01	-0,10	-3,00	0,00	0,91	1,10
PesoEx	0,02	0,01	0,09	2,67	0,01	0,94	1,07
PesoP	0,02	0,01	0,07	2,15	0,03	0,97	1,03

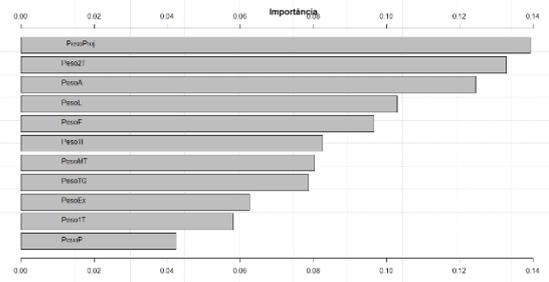
Anexo C – Figuras dos gráficos de importância do Estudo 1 de EDM



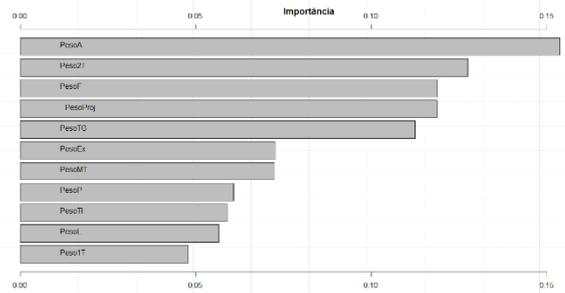
CV



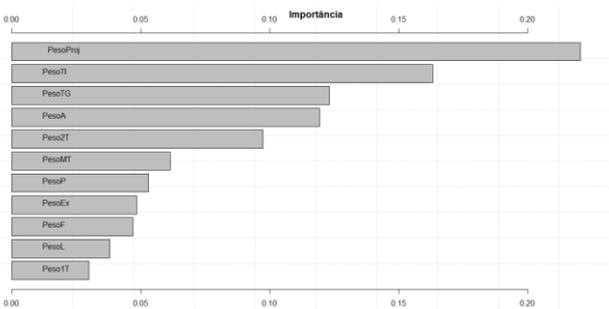
MARS



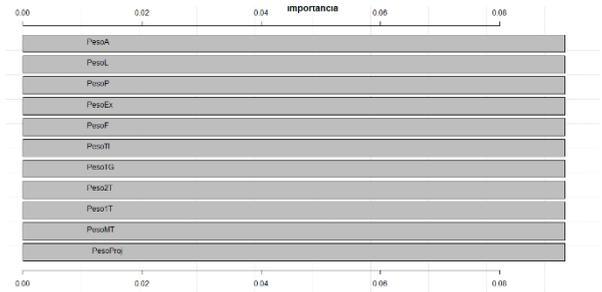
MLP



MLPE

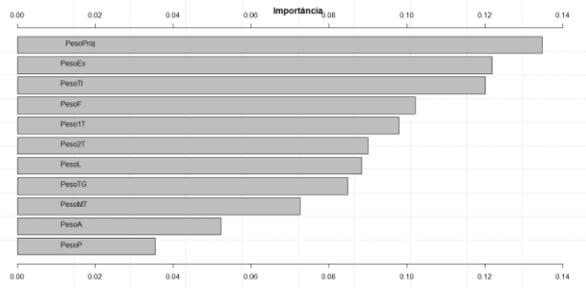


MR

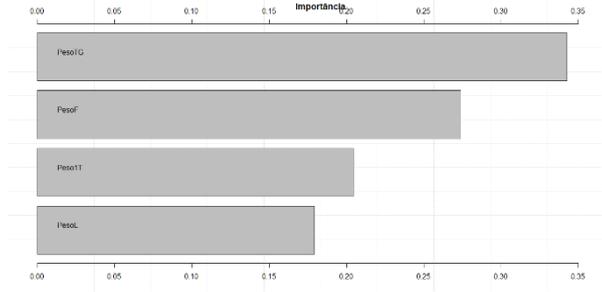


RPART

A influência dos métodos de avaliação nos resultados dos alunos do ensino superior

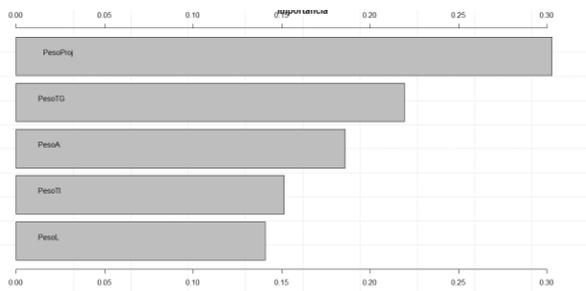


RVM

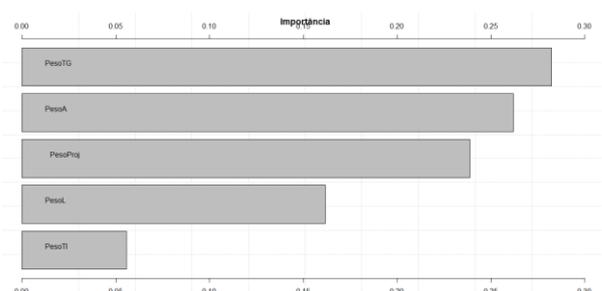


XGBOOST

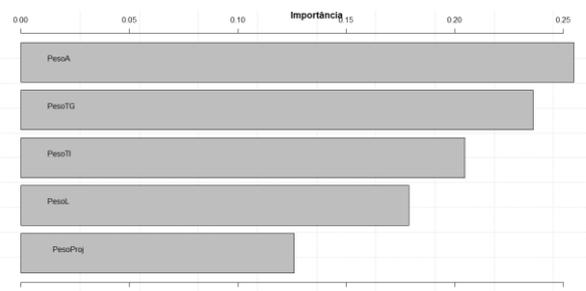
Anexo D – Figuras dos gráficos de importância do Estudo 2 de EDM



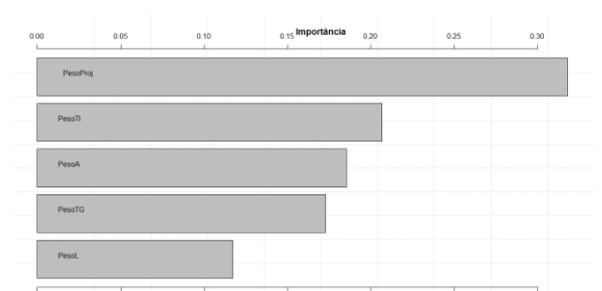
CV



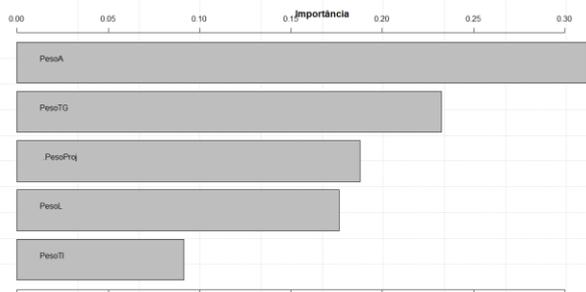
MARS



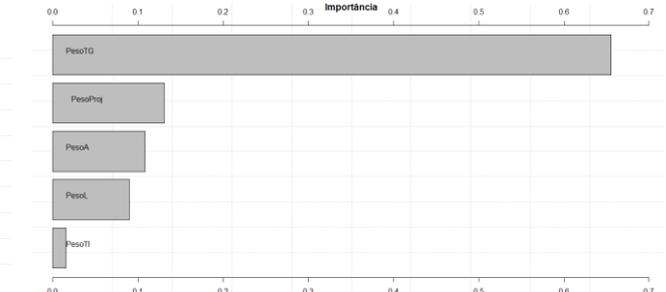
MLP



MR

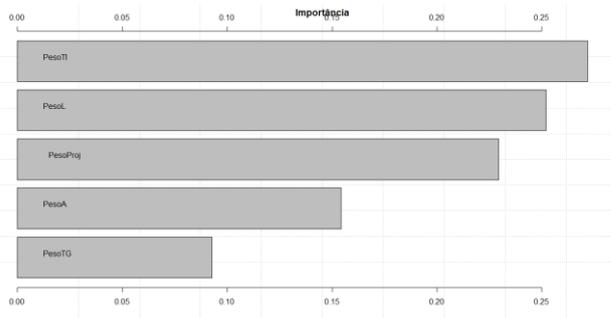


RANDOMFOREST

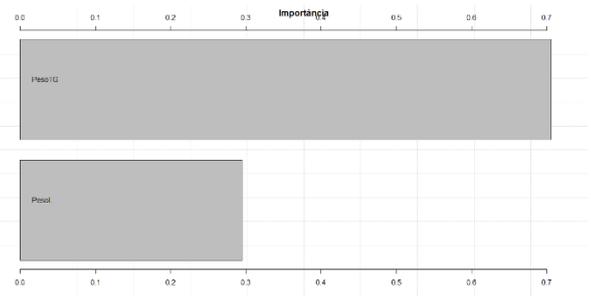


RPART

A influência dos métodos de avaliação nos resultados dos alunos do ensino superior

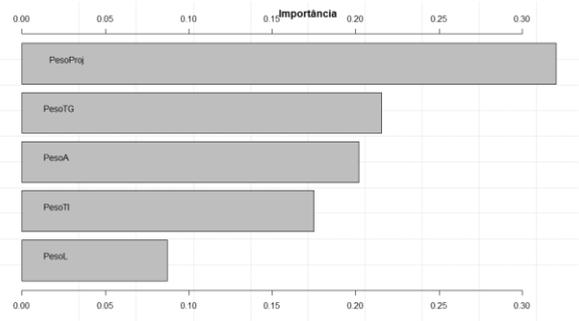


RVM

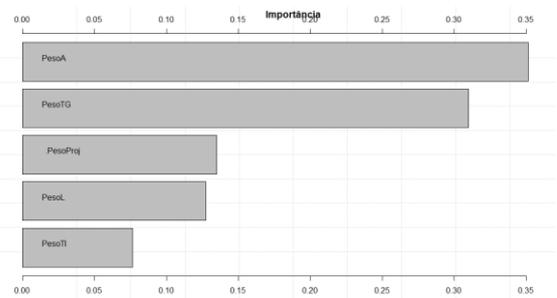


XGBOOST

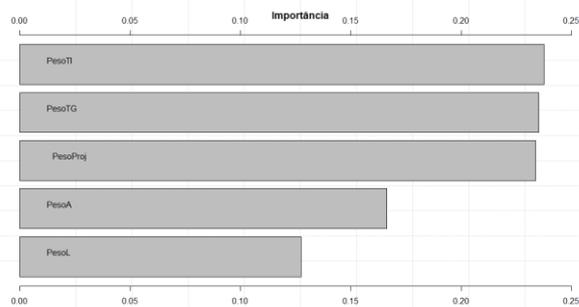
Anexo E – Figuras dos gráficos de importância do Estudo 3 de EDM



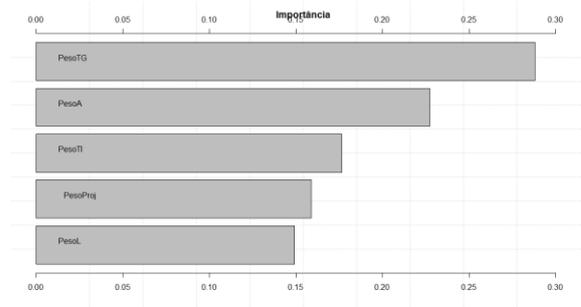
CV



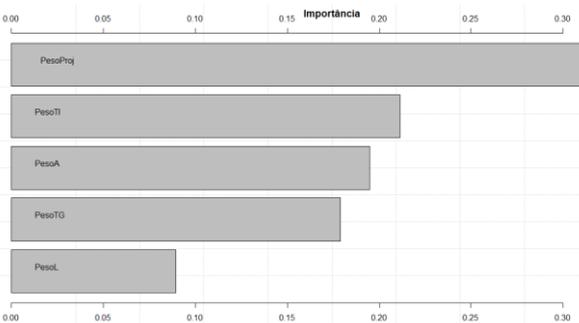
MARS



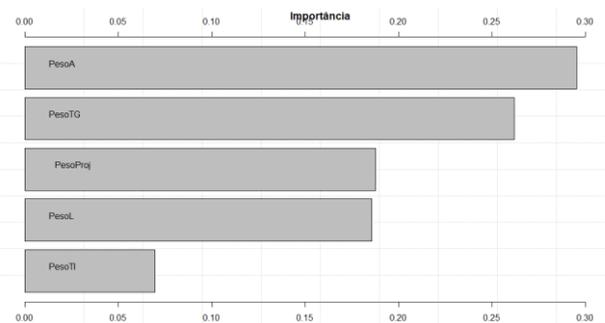
MLP



MLPE

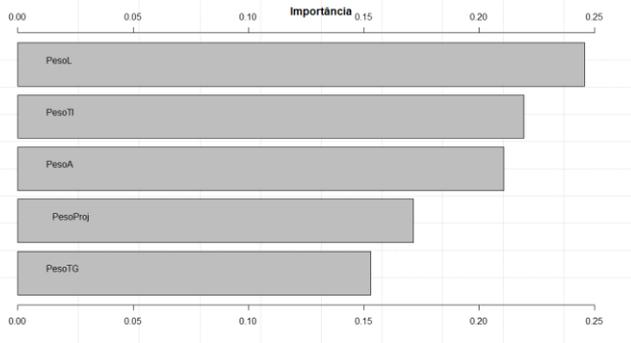


MR

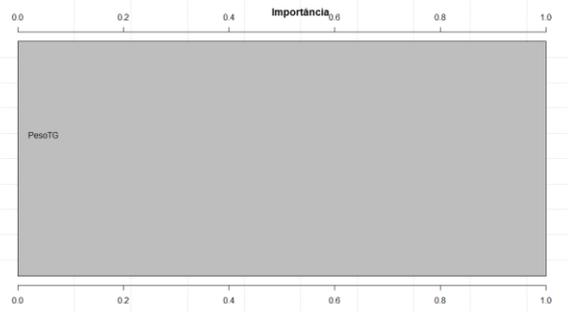


RANDOMFOREST

A influência dos métodos de avaliação nos resultados dos alunos do ensino superior

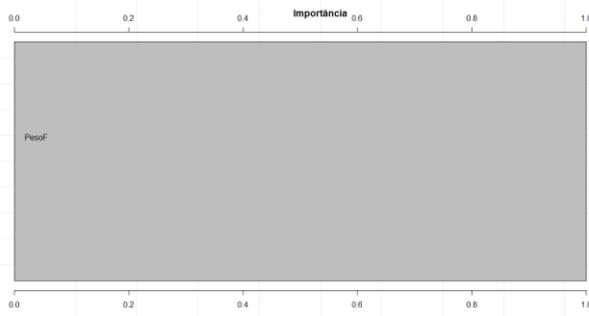


RVM

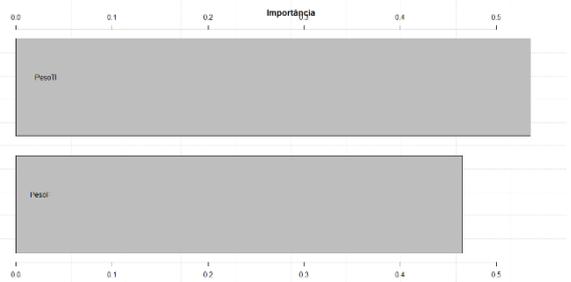


XGBOOST

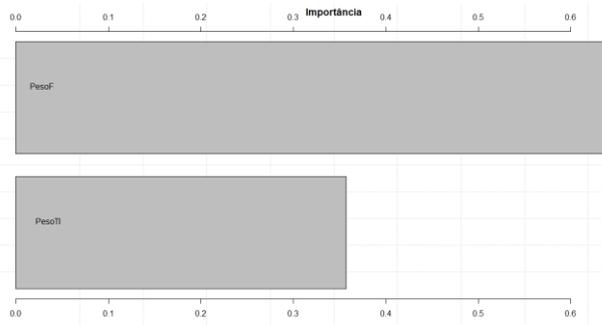
Anexo F – Figuras dos gráficos de importância do Estudo 4 de EDM



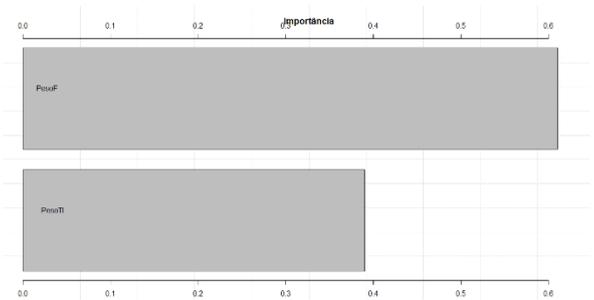
CV



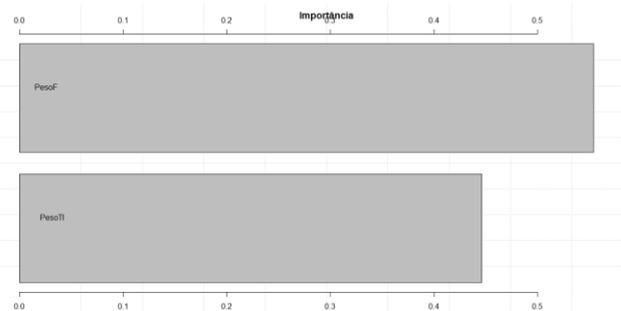
MARS



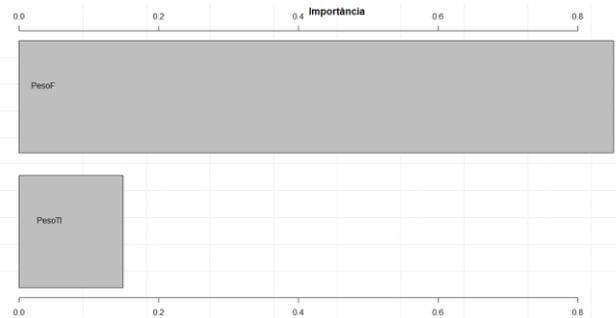
MLP



MLPE

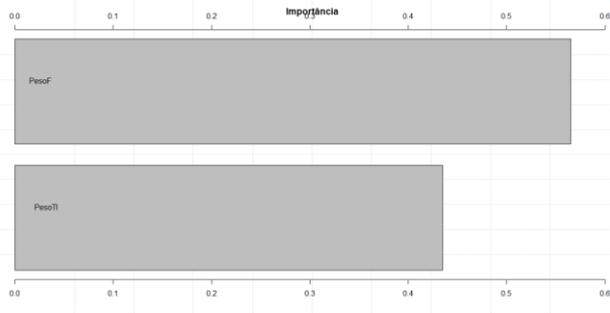


MR

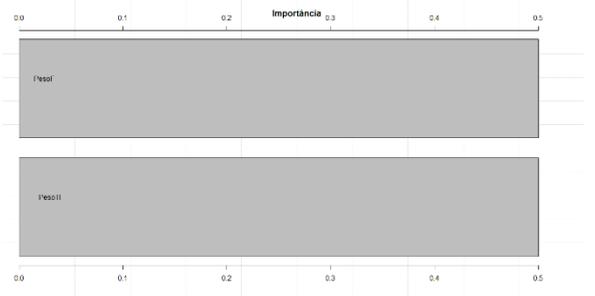


RPART

A influência dos métodos de avaliação nos resultados dos alunos do ensino superior



RVM



XGBOOST