Detecção de Traços de Personalidade em Textos para Apoiar a Formação de Grupos para Colaboração

Taís Borges Ferreira¹, Márcia Aparecida Fernandes¹

¹Faculdade de Computação – Universidade Federal de Uberlândia (UFU)
Uberlândia – MG – Brasil

{taisbferreira,marcia}@ufu.br

Abstract. *In order to provide a way to detect personality traits transparently and to make this information available to support grouping strategies for collaboration, we evaluate texts written by students that can be captured, for instance, in virtual learning environments. The results of text evaluation indicate as more effective to perform this task the algorithms M5P, SMO and LWL. Aiming to understand how students organize themselves to reach a goal, we observed students working in groups to build a base of groups and their performance to support grouping suggestions based on personality traits.*

Resumo. *Com o objetivo de fornecer um meio de detectar traços de personalidade de forma transparente e disponibilizar esta informação para apoiar estratégias de agrupamento para colaboração, avaliamos textos escritos por alunos que podem ser capturados, por exemplo, em ambientes virtuais de aprendizagem. Os resultados da avaliação de textos indicam como mais eficaz para inferir os traços de personalidade dos algoritmos M5P, SMO e LWL. Também avaliamos estudantes trabalhando em grupo para entender como eles se organizam para atingir uma meta e formar uma base de combinações de alunos para apoiar o agrupamento usando os traços de personalidade.*

1. Introdução

A aprendizagem colaborativa envolve a construção de conhecimento por meio da interação entre os membros de um grupo para atingir uma meta compartilhada [Hron e Friedrich 2003]. A efetividade da colaboração para a aprendizagem apoia-se nas teorias que definem o aprendizado como um processo social que ocorre de forma mais efetiva por meio de interações interpessoais em um contexto cooperativo [Vygotsky 1978]. Assim, um indivíduo trabalhando em colaboração com outros e em situações que envolvam resolução de problemas, é exposto a outros pontos de vista, o que o leva a questionar seu entendimento inicial e motiva o seu aprendizado [Alavi 1994].

Como a formação do grupo influencia a maneira que as pessoas trabalham juntas para atingir um objetivo, uma forma de dar suporte à colaboração é preocupar-se com a formação dos grupos, o que é considerado crucial para desencadear interações produtivas entre os pares [Magnisalis et al. 2011, Manske et al. 2015]. O agrupamento de forma descuidada pode, por exemplo, ser o gatilho para situações indesejáveis, como o isolamento de algum aluno dentro de um grupo. Por esse motivo, é essencial considerar uma estratégia de formação de grupos para promover interações produtivas e garantir que tal estratégia possa ser usada para dar apoio à colaboração.

Conhecer os estados afetivos dos estudantes, como emoções e traços de personalidade, permite pensar uma estratégia de formação de grupos de estudantes que evite combinações indesejáveis para a colaboração. Por exemplo, ao encontrar um traço relacionado ao desempenho acadêmico, como conscienciosidade [Papamitsiou e Economides 2014], usá-lo na escolha dos indivíduos pode ser pensado como uma forma de tentar balancear os grupos em relação ao conhecimento e desempenho acadêmico. Os traços de personalidade também apontam tendências de comportamento que podem influenciar as interações entre os alunos em um grupo. Por exemplo, para funcionar de maneira efetiva, um grupo não deveria ser formado por vários indivíduos com alto grau de extroversão, uma vez que eles poderiam se distrair com a interação e negligenciar a meta comum [Altanopoulou e Tselios 2015].

Dentre as várias construções psicométricas encontradas na literatura para medir aspectos de personalidade, a mais usada é o Five Factor Model [Altanopoulou e Tselios 2015]. No entanto, usar questionários para avaliar alunos é indesejável por diversos motivos como aumento da carga cognitiva do estudante. Assim, extrair tais características a partir do texto que o aluno produz de suas atividades usuais nos chats e fóruns de plataformas de ensino ajuda a conhecê-lo sem envolvê-lo diretamente na avaliação. Como [Wen et al. 2014] sugere, os chats e fóruns são fontes de informação relevante para aprender sobre a motivação e a atividade de um aluno em um ambiente virtual de aprendizagem (AVA). Além disso, como as pessoas usam a linguagem para expressar seus pensamentos e emoções, ela geralmente carrega as características do interlocutor e, portanto, é possível encontrar uma relação entre as palavras escolhidas por um aluno em uma conversa e sua personalidade [Tausczik e Pennebaker 2010].

Com o objetivo de fornecer um meio de detectar traços de personalidade de forma transparente e tornar essa informação disponível para apoiar estratégias de formação de grupos para colaboração, este trabalho trata da avaliação de textos escritos por alunos, que podem, por exemplo, ser capturados em chats e fóruns em um AVA. O conhecimento sobre os traços de personalidade do aluno extraídos do texto serão usados como critério de formação de grupos. Para observar a influência dos traços de personalidade na interação dos alunos em grupo, observamos também duas turmas de alunos trabalhando de forma colaborativa. O artigo está organizado conforme se segue. A segunda seção apresenta e discute trabalhos relacionados. A descrição do modelo para formação de grupos se encontra na terceira seção. Os detalhes dos experimentos realizados e análise dos resultados estão na quarta seção. Conclusão e trabalhos futuros são apresentados na quinta seção.

2. Revisão da literatura

Os fundamentos teóricos desta pesquisa incluem teorias sobre traços de personalidade e formação de grupos (também dito agrupamento) para aprendizagem colaborativa. Alguns trabalhos que serão apresentados a seguir unem estes dois temas, pois sugestão de agrupamentos é uma das aplicações da teoria de traços de personalidade, sendo Five Factor Model, também conhecido como Big Five, um expoente neste segmento.

Big Five é uma organização hierárquica de traços de personalidade em cinco dimensões básicas que agrupam várias características e é considerado confiável na avaliação de traços de personalidade por ter sido amplamente estudado e testado em diversas culturas [John e Srivastava 1999, Goldberg 1981]. As cinco dimensões avaliadas com os in-

ventários do Big Five são extroversão, relacionada à tendência de buscar atividades sociais e/ou excitantes, neuroticismo, que relaciona-se à baixa tolerância ao estresse, conscienciosidade, que é relacionada ao quão cuidadoso, dedicado e comprometido um indivíduo é, afabilidade, que diz respeito à tendência de ser cooperativo e dar atenção às questões do outro, e abertura, relacionada ao gosto por novas experiências.

Os comportamentos esperados para cada traço de personalidade podem guiar o entendimento de como os indivíduos atuam em grupo e ajudar na seleção de alunos para formar bons grupos para colaboração. [Roberts et al. 2008], por exemplo, apontam indivíduos extrovertidos como mais eficientes na construção de novas conexões e por isso podem atuar como aqueles que estabelecem a ligação com os demais membros de um grupo. [Spoelstra et al. 2015] propõe o agrupamento de indivíduos conscienciosos, além de usar suas preferências e conhecimentos, para formar grupos produtivos e de aprendizagem, dada a característica de tais indivíduos serem organizados e orientados à meta.

Apesar de encontrar relação positiva entre conscienciosidade e o ganho de aprendizagem dos alunos trabalhando em uma atividade colaborativa, [Altanopoulou e Tselios 2015] propõe a distribuição dos indivíduos com alto grau de conscienciosidade entre os grupos para que eles possam influenciar a habilidade do grupo de cumprir o prazo proposto. Seus resultados sugerem ainda que grupos formados por indivíduos com alto grau de extroversão são ruins para a aprendizagem pois tais indivíduos tendem a se distrair mais fácil com a interação social.

[Akhtar et al. 2015] apontam a abertura como o segundo preditor mais forte para o engajamento no trabalho e atribuem esse resultado a capacidade de indivíduos com alto grau de abertura lidarem de forma eficiente com o seu ambiente. Os resultados de [Bozionelos 2017] apontam alta conscienciosidade como causa de negligência das relações sociais em função da preocupação com as atividades e alto grau de neuroticismo como benéfico para um grupo, uma vez que é um traço de personalidade ligado à preocupação, o que pode fazer um indivíduo agir de forma a manter os laços sociais já estabelecidos, ainda que sejam indivíduos que evitem iniciar a interação.

De acordo com os resultados obtidos nos estudos citados, o agrupamento de alunos para uma atividade colaborativa, deveria considerar a escolha de pelo menos um indivíduo consciencioso, para garantir que alguém vai se preocupar com a meta, pelo menos um indivíduo extrovertido para iniciar conexões entre os membros de um grupo, evitar agrupamentos de alunos muito extrovertidos, para evitar distração em função da interação social, e evitar grupos com todos os indivíduos conscienciosos, para evitar que as interações sociais sejam negligenciadas e impeçam a troca de conhecimento.

Mesmo a personalidade sendo reconhecida como fator importante na construção de sistemas de recomendação personalizados, o esforço de capturá-la de forma explícita é indesejado e pode atrapalhar seu uso para recomendações no mundo real [Wu e Chen 2015]. Em um AVA, onde o aluno não tem a supervisão presencial de um tutor, torna-se difícil conhecer suas características afetivas e usá-las para apoiar o processo de aprendizagem. Neste sentido, vários trabalhos foram desenvolvidos para extrair conhecimento sobre o aprendiz de tais plataformas.

[Wen et al. 2014] conduziram análises de linguística aplicadas às postagens de alunos nos fóruns para encontrar a relação entre os textos escritos e a motivação e

o envolvimento cognitivo do aluno no contexto de cursos on-line maciços (MOOC's). [Wen et al. 2014] enfatizam que o uso de análises baseadas em características linguísticas definidas na literatura, por exemplo, categorias de palavras, é efetivo na prática e de mais fácil generalização para outros domínios do que modelos de palavras como saco de palavras (bag-of-words). De forma semelhante ao seu estudo, a estratégia proposta nesta pesquisa envolve o uso das categorias de palavras do dicionário LIWC.

A linguagem é usada para expressar pensamentos e emoções, por isso, carrega parte das características do interlocutor e, portanto, além das informações sobre motivação dos aprendizes é possível extrair dos textos que um aluno escreveu seus traços de personalidade [Tausczik e Pennebaker 2010]. [Machado et al. 2015], por exemplo, conduziram um estudo inicial com objetivo de construir um léxico para ajudar na extração de informações relativas à teoria do Big Five de textos escritos em Português (que, neste texto, se refere ao idioma Português do Brasil) e encontrou algumas relacionadas aos traços. [Paim et al. 2016] desenvolveu um método de inferência da personalidade por meio de textos em Português, adquiridos na rede social Facebook, usando algoritmos de regressão para construir seus modelos de inferência.

Tanto [Machado et al. 2015] quanto [Paim et al. 2016] indicam possibilidade de contribuição para melhorar a qualidade dos métodos de inferência da personalidade por meio de textos escritos em Português. Em relação aos algoritmos usados na construção de modelos de detecção automática de características de personalidade a partir do texto, o mais comum é o uso de algoritmos indutores de regressão como [Wu e Chen 2015] e [Paim et al. 2016]. Além dos algoritmos escolhidos por ambos, também foram incluídos nos experimentos outros algoritmos, especificados na subseção 4.2, com o objetivo de selecionar um algoritmo de aprendizagem de máquina adequado para determinar cada um dos traços de personalidade a partir dos textos escritos pelos estudantes.

3. Agrupamento utilizando Traços de Personalidade

Os trabalhos discutidos na seção anterior mostram que um aspecto importante da aprendizagem colaborativa é a formação de grupos e que esta pode ser conduzida observando, por exemplo, características como os traços de personalidade dos estudantes. Assim, neste trabalho considera-se o uso da teoria do Big Five para identificação dos traços, a partir de textos escritos por estudantes, que são obtidos das interações dos estudantes em um AVA. A viabilidade de obtenção destes textos pode ser constatada no fato de que a maioria dos AVAs possuem ferramentas de comunicação, tais como fóruns e chats, utilizadas em atividades propostas aos estudantes. No Moodle, por exemplo, o conteúdo das postagens nos fóruns e mensagens trocadas entre os usuários, bem como o autor de cada texto, são gravadas em um banco de dados relacional. A partir de uma consulta nesse banco, é possível obter amostras de textos escritos pelos alunos.

A proposta consiste então em analisar os textos extraídos do AVA, identificar os traços de personalidade e propor formação de grupos, observando os traços e visando a configuração mais adequada ao processo de aprendizagem. A Figura 1 descreve este modelo, onde o *Banco de Dados* fornece as informações relativas aos textos escritos. No módulo de formação de grupos da Figura 1, os textos são processados usando biblioteca NLTK do Python [Bird et al. 2009], criada para processamento de linguagem natural. Na etapa *Tratamento de texto*, stopwords, acentuação e pontuação são remo-

Figura 1. Formação de grupos com traços de personalidade

vidas e uma lista de palavras é extraída. Cada palavra desta lista é associada a categorias, usando o dicionário interno do Linguistic Inquiry and Word Count (LIWC) [Tausczik e Pennebaker 2010]. O dicionário possui categorias de palavras relativas a processos psicológicos e cognitivos, como raiva, tristeza e insight, que podem ser relacionadas aos traços de personalidade do Big-Five [Neto et al. 2017]. A versão do LIWC escolhida para a categorização de palavras foi o Brazilian Portuguese LIWC 2007 Dictionary [Balage Filho et al. 2013].

A frequência das categorias no texto processado é utilizada pelo *Classificador dos traços Big Five*. A classificação do aluno, dada pelos percentuais de cada traço calculados a partir do texto, é usada na etapa *Sugestão de grupos*, onde as combinações de traços de personalidade são analisadas, levando em consideração informações decorrentes de *Avaliação dos grupos*. A avaliação dos grupos ocorre por meio das informações extraídas dos logs das atividades de um grupo resolvendo uma tarefa colaborativa na etapa de *Extração de métricas*. São considerados bons grupos aqueles que conseguem terminar a atividade no prazo e, além disso, todos contribuem para a solução respondendo aos exercícios e revisando as respostas dos colegas. Os grupos e seus respectivos desempenhos, calculados na avaliação dos grupos, são incluídos em uma base de combinações de indivíduos usada para formar novos grupos usando os traços de personalidade.

Como o fluxo deste modelo funciona de maneira dinâmica e cíclica, as interações entre estudantes, por meio das ferramentas de comunicação do AVA, devem gerar mais trocas de mensagens e, conseqüentemente, mais textos serão coletados, permitindo ajustes na classificação de cada um dos traços de personalidade. Também, cada vez que novos grupos trabalham em uma atividade colaborativa, a base usada na sugestão de grupos pode ser revisada e atualizada.

4. Experimentos e Análise de Resultados

Dois experimentos foram conduzidos com objetivo de desenvolver a proposta de formação de grupos usando traços de personalidade deduzidos a partir do texto. Nesta seção são apresentados os resultados obtidos nesses experimentos a partir dos dados coletados.

4.1. Coleta de Dados e Tratamento dos Textos

Inicialmente, a coleta de textos escritos em Português e dos escores dos participantes em cada um dos cinco traços de personalidade do Big-Five foi feita por meio de um questionário. As questões da primeira parte do questionário são as perguntas do inventário

de 44 itens do Big-Five, que avalia os cinco traços de personalidade de forma eficiente e flexível [John e Srivastava 1999]. A versão do questionário usada neste trabalho foi traduzida para Português e sua validade testada no Brasil [Andrade 2008]. A coleta das respostas nesta parte do questionário ocorreu por meio do uso de uma escala likert de 5 pontos. As demais perguntas do questionário são cenários nos quais o aluno teve que descrever sua possível resposta ou reação em cada um deles. Esses cenários foram acrescentados com o objetivo capturar textos escritos para montar a base de treinamento dos modelos de inferência dos traços de personalidade. Os participantes foram orientados a pensar em sua interação com seus amigos ou colegas de turma para responder às perguntas de forma a permitir capturar linguagem próxima àquela que usa em suas conversas.

Os alunos convidados a responder o questionário estavam matriculados nas disciplinas de Análise Financeira, Finanças de Curto Prazo e Dados e Informações Empresariais, da Faculdade de Gestão e Negócios da Universidade Federal de Uberlândia e os alunos de ensino médio técnico matriculados nas disciplinas de Engenharia de Software e Lógica de Programação do Instituto Federal Goiano - Campus Campos Belos. A resposta ao questionário foi facultativa. Um total de 135 alunos, com idade entre 15 e 46 anos, responderam ao questionário: 76 do sexo feminino e 59 do sexo masculino. Destes, 29 alunos de ensino médio técnico em Informática e 106 de ensino superior em Administração e Gestão da Informação.

4.2. Determinando Traços de Personalidade

A base de 135 amostras coletadas com o questionário foi dividida de forma aleatória em 80% treinamento e 20% teste. Dentre os algoritmos de aprendizagem de máquina disponíveis na ferramenta Weka, foram treinados para classificar os traços de personalidade abertura (O), conscienciosidade (C), extroversão (E), afabilidade (A) e neuroticismo (N) os algoritmos de regressão linear (RL), otimização sequencial mínima (SMO), aprendizagem localmente ponderada (LWL), função de base radial (RBF), processo gaussiano (GP), as árvores de decisão M5P e REPTree e rede perceptron multicamadas (MLP). O erro médio quadrático, calculado para avaliar a qualidade da predição de cada algoritmo em relação aos valores obtidos com o questionário, é apresentado na Tabela 1.

Tabela 1. Erro quadrático médio por traços de personalidade.

	GP	LWL	M5P	MLP	RBF	RL	REPTree	SMO
O	207,19	193,70	150,40	1.245,30	770,33	1.651,32	218,62	323,14
C	158,02	179,28	227,64	524,74	380,48	301,83	219,47	132,75
E	369,78	319,35	383,78	1.277,03	1.856,88	923,17	487,25	457,84
A	192,54	179,41	182,46	488,86	270,91	215,48	190,50	262,46
N	400,09	413,55	349,77	1.810,89	825,12	541,25	726,11	424,07

Quanto menor o valor do erro calculado, maior a chance dos valores previstos pelo algoritmos se aproximarem do escore obtido com o questionário. Para os cinco traços, os algoritmos MLP e RBF tiveram desempenho insuficiente em relação ao erro. Os menores erros foram obtidos com M5P para neuroticismo e abertura, SMO para conscienciosidade e LWL para extroversão e afabilidade. Apesar do menor erro para afabilidade ter sido obtido com LWL, os valores previstos pelo algoritmo foram praticamente todos em torno

de 63%, independente do valor dos atributos de entrada nas amostras de teste. Neste caso, M5P seria uma opção melhor que o LWL.

Para avaliar os algoritmos quanto à quantidade de valores previstos (VP) próximos aos valores do questionário (VQ), os gráficos da Figura 2 foram construídos contando o número de indivíduos cujo valor de $|VQ - VP| \leq 10$. Por exemplo, um indivíduo com $VQ = 60$ e $50 \leq VP \leq 70$ é contabilizado como menor que 10. Os algoritmos que conseguiram classificar um número maior de indivíduos dentro desse intervalo foram M5P para abertura, M5P e SMO para afabilidade e neuroticismo, SMO para conscienciosidade e LWL para extroversão.

De acordo com os resultados, os algoritmos M5P, LWL e SMO apresentaram os melhores valores para a classificação dos cinco traços de personalidade. O algoritmo GP, apesar de não ser o melhor, teve resultados próximos ao melhor e, portanto, pode ser considerado como outra opção para inferência dos traços.

Com todos os algoritmos, extroversão e neuroticismo pareceram mais difíceis de classificar, apresentando erros maiores em comparação aos demais traços e classificando a maioria dos indivíduos com $|VQ - VP| > 10$. Um motivo poderia ser o tamanho dos textos avaliados que tinham em média 51 palavras, o que pode ser considerado um volume baixo para treinamento. Uma opção para melhorar a qualidade da predição é capturar um volume maior de textos dos alunos para gerar a base de treinamento.

Figura 2. Contagem de indivíduos classificados com $|VQ - VP| \leq 10$

4.3. Interação nas Atividades em Grupo

Em duas turmas que responderam ao questionário, os alunos foram observados trabalhando em grupo. Na turma 1, os grupos foram formados com 3 alunos. A atividade foi concluída em horário de aula e consistia de 6 questões, propostas pelo professor, para reforçar e testar os conhecimentos na disciplina de Engenharia de Software. A nota de cada grupo foi atribuída pelo professor da turma após análise e avaliação das respostas dadas às questões. Na turma 2, os grupos formados tinham de 3 a 5 alunos que tiveram uma semana para concluir a tarefa. A atividade envolvia 12 questões relacionadas à aplicação prática dos conhecimentos da disciplina de Análise Financeira, como cálculo de balanço patrimonial. Todos que entregaram a tarefa receberam a mesma nota.

Em ambas as turmas, os alunos ficaram livres para escolher a composição dos grupos. A participação foi compulsória e a atividade realizada era parte da avaliação da disciplina. As atividades tiveram que ser concluídas usando a ferramenta de escrita compartilhada do Google Drive. A qualidade dos grupos foi avaliada a partir do histórico de edições dos alunos registradas pela ferramenta. Um grupo foi considerado bom se: 1) a atividade foi concluída, 2) todos os membros fizeram alterações no documento, 3) algum membro alterou a resposta de outro e 4) todos os membros fizeram contribuições significativas.

Na turma 1, exceto pelo grupo que recebeu a pior nota, todos participaram da solução e concluíram a tarefa, respondendo ou revisando as respostas dos outros. Na turma 2, nem todos os grupos reuniram esforços para completar a atividade. Apenas o grupo de 3 membros completou a atividade e, além disso, todos os membros interagiram e deram contribuição significativa. Nos demais grupos, com mais de 3 alunos, pelo menos um indivíduo não participou da tarefa. Esse resultado sugere que os grupos deveriam ser formados por, no máximo, 3 membros, para que todos participem da solução.

A Tabela 2 apresenta os escores dos alunos da turma 1 nos traços de personalidade abertura (O), conscienciosidade (C), extroversão (E), afabilidade (A) e neuroticismo (N). O símbolo B representa escores menores que 38%, M representa escores entre 38% e 62% e A, maiores que 62%. A nota atribuída pelo professor, apresentada na Tabela 2, corresponde a um valor entre 0 e 3. Dois comportamentos foram observados nos grupos: mais de um membro do grupo editou a mesma questão ou a divisão clara de quem respondeu e quem revisou cada exercício. O comportamento de editar juntos ocorreu apenas nos grupos 2, 4, 6 e 8 da turma 1, nos quais a maioria dos membros apresentou alto grau de abertura.

Tabela 2. Escores nos traços de personalidade e nota dos alunos da turma 1.

Grupo	aluno 1					aluno 2					aluno 3					Nota
	O	C	E	A	N	O	C	E	A	N	O	C	E	A	N	
5	A	M	M	M	B	M	M	A	A	B	M	M	M	A	B	2,5
2	A	M	B	A	A	A	M	A	A	A	A	M	M	A	M	2,2
1	A	A	M	A	A	M	M	M	M	A	M	A	M	A	B	2
7	M	M	A	A	B	A	A	A	M	M	M	A	M	M	B	1,8
4	A	A	A	A	M	A	A	A	A	M	A	A	A	A	B	1,6
6	A	M	A	A	A	A	M	M	B	A						1,5
8	A	M	M	M	A	A	M	M	M	A	M	M	B	A	B	1,5
3	A	A	M	A	M	M	M	M	A	M	M	A	M	M	M	1,1

Ter indivíduos conscienciosos pareceu aumentar a chance de ter alguém envolvido na solução da atividade, já que nos grupos da turma 2 foi possível observar que os membros mais conscienciosos assumem a responsabilidade de fazer a atividade. Por outro lado, na turma 1, os grupos 2 e 5, que receberam maior nota do professor, não tinham membros com alta conscienciosidade, enquanto a pior nota foi do grupo 3, com dois indivíduos conscienciosos. Além disso, um membro do grupo 3 não se envolveu na atividade. Portanto, apesar de aumentar a chance de entrega da atividade, ter indivíduos conscienciosos entre os membros do grupo não garantiu seu bom funcionamento.

Além da presença de membros com traços de personalidade considerados preditores do desempenho do grupo, notou-se que as características dos demais membros são relevantes para o desempenho. Os resultados encontrados sugerem que a combinação de traços altera o grau de influência de um determinado traço no grupo, como observado em relação à conscienciosidade. Portanto, formar grupos usando uma base de bons grupos parece melhor que a seleção de membros avaliando apenas seus escores nos traços. Por isso, as observações resultantes dos experimentos foram empregadas na formação da base inicial para apoiar a sugestão de novos grupos, e que deve ser atualizada e aprimorada com experimentos futuros.

5. Conclusão e Trabalhos Futuros

Com o objetivo de fornecer um meio de detectar traços de personalidade, de forma transparente, e disponibilizar essa informação para apoiar estratégias de agrupamento para colaboração, avaliamos textos escritos por alunos. Os resultados da avaliação dos textos indicam os algoritmos M5P, SMO e LWL como mais efetivos para inferir os traços de personalidade. A partir da classificação dos alunos e com base nas avaliações das atividades em grupo, a base inicial de combinações que resultaram em bons grupos foi criada.

Os próximos passos no desenvolvimento deste trabalho é melhorar a qualidade da classificação dos modelos de previsão de traços de personalidade. Como a base de textos utilizada tinha em média 51 palavras e os algoritmos testados tiveram dificuldade para prever corretamente neuroticismo e extroversão, um caminho é repetir os experimentos com um volume maior de texto. Em relação à sugestão de grupos, a base inicial será atualizada e ampliada com novos experimentos, observando os grupos formados a partir das combinações encontradas nos resultados.

Referências

- Akhtar, R., Boustani, L., Tsivrikos, D., e Chamorro-Premuzi, T. (2015). The engageable personality: Personality and trait EI as predictors of work engagement. *Personality and Individual Differences*, 73:44–49.
- Alavi, M. (1994). Computer-mediated collaborative learning: An empirical evaluation. *Journal MIS Quarterly*, 18:159–174.
- Altanopoulou, P. e Tselios, N. (2015). How does personality affect wiki-mediated learning? In *Proceedings of International Conference on Interactive Mobile and Communication Technologies and Learning*, pages 16–18.
- Andrade, J. M. (2008). *Evidências de Validade do Inventário dos Cinco Grandes Fatores de Personalidade para o Brasil*. PhD thesis, Instituto de Psicologia - Universidade de Brasília.
- Balage Filho, P. P., Aluísio, S. M., e Pardo, T. A. S. (2013). An evaluation of the brazilian portuguese liwc dictionary for sentiment analysis. In *Proceedings of Brazilian Symposium in Information and Human Language Technology*, pages 215–219.
- Bozionelos, G. (2017). The relationship of the big-five with workplace network resources: More quadratic than linear. *Personality and Individual Differences*, 104:374–378.
- Goldberg, L. R. (1981). Language and individual differences: The search for universal in personality lexicons. *Review of personality and social psychology*, 2:141–166.

- Hron, A. e Friedrich, H. F. (2003). A review of web-based collaborative learning: factors beyond technology. *Journal of Computer Assisted Learning*, 19:70–79.
- John, O. e Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2:102–138.
- Machado, A. A., Longhi, M. T., Nunes, M. A. S. N., e Pardo, T. A. S. (2015). Personalidade lexicon: Um léxico em português brasileiro para mineração de traços de personalidade em textos. In *Anais do XXVI Simpósio Brasileiro de Informática na Educação*, pages 1122–1126.
- Magnisalis, I., Demetriadis, S., e Karakostas, A. (2011). Adaptive and intelligent systems for collaboration learning support: A review of the field. *IEEE Transactions on Learning Technologies*, 4:5–20.
- Manske, S., Hecking, T., Chounta, I., e Hoppe, H. (2015). Using differences to make a difference: a study on heterogeneity of learning groups. In *Proceedings of International Conference on Computer Supported Collaborative Learning*, pages 182–189.
- Neto, A. T., Ferreira, T., e Fernandes, M. (2017). Development of a big-five personality traits classification approach via analysis of texts in brazilian portuguese. In *Anais do IV WICSI - XIII Simpósio Brasileiro de Sistema de Informação*, pages 57–60.
- Paim, A., Camati, R., e Enembreck, F. (2016). Inferência de personalidade a partir de textos em português utilizando léxico linguístico e aprendizagem de máquina. In *Anais do XIII Encontro Nacional de Inteligência Artificial e Computacional*, pages 481–492.
- Papamitsiou, Z. e Economides, A. A. (2014). The effect of personality traits on students' performance during computer-based testing: a study of the big five inventory with temporal learning analytics. In *Proceedings of IEEE 14th International Conference on Advanced Learning Technologies*, pages 378–382.
- Roberts, S. G. B., Wilson, R., Fedurek, P., e Dunbar, R. I. M. (2008). Individual differences and personal social network size and structure. *Personality and Individual Differences*, 44:954–964.
- Spoelstra, H., Van Rosmalen, P., Houtmans, T., e Sloep, P. (2015). Team formation instruments to enhance learner interactions in open learning environments. *Computers in Human Behavior*, 45:11–20.
- Tausczik, Y. R. e Pennebaker, J. W. (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29:24–54.
- Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.
- Wen, M., Yang, D., e Rose, C. P. (2014). Linguistic reflections of student engagement in massive open online courses. In *Proceedings of the International Conference on Weblogs and Social Media*.
- Wu, W. e Chen, L. (2015). Implicit acquisition of user personality for augmenting movie recommendations. In *Proceedings of International Conference on User Modeling, Adaptation and Personalization*, pages 302–314.