

Detecção automática de traços de personalidade e recomendação de agrupamento com o modelo Big Five

Taís Borges Ferreira¹, José Antonio Buiar², Márcia Aparecida Fernandes¹,
Andrey Ricardo Pimentel³, Luiz Eduardo S. Oliveira³

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

²Universidade Tecnológica Federal do Paraná (UTFPR)
Curitiba – PR – Brasil

³Universidade Federal do Paraná (UFPR)
Curitiba – PR – Brasil

taisbferreira@ufu.br, buiar@utfpr.edu.br, marcia@ufu.br
{andrey, luiz.oliveira}@inf.ufpr.br

Abstract. *The learning experience in virtual environments can be significantly improved through the effective introduction of collaborative learning. The group formation of students is a crucial aspect in a such learning. Due to the lack of interaction between students, this task becomes complex, requiring automatic and intelligent tools to help students, tutors and environments to determine groupings for collaborative work. This article is a joint effort of two research groups, aimed at fostering the advancement of collaborative learning, presenting the use of personality traits of the Big Five model, identified through natural language texts produced by students. The experiments performed independently were compared, demonstrating the feasibility of the proposals.*

Resumo. *A experiência de aprendizagem em ambientes virtuais pode ser significativamente melhorada através da introdução efetiva de aprendizagem colaborativa. A formação de grupos é um aspecto crucial para este tipo de aprendizagem. Devido à ausência de interação entre estudantes, esta tarefa torna-se complexa, sendo necessárias ferramentas para auxiliar a determinar agrupamentos para trabalho colaborativo. Este artigo é um esforço conjunto de dois grupos de pesquisa, visando promover o avanço da aprendizagem colaborativa, utilizando o modelo Big Five de personalidade, identificados nos textos produzidos pelos alunos. Os experimentos realizados independentemente foram comparados, demonstrando a viabilidade das propostas.*

1. Introdução

A ausência de interação física entre estudantes em ambientes virtuais de aprendizagem tem sido considerado um grande desafio [Kop 2011, Everson 2011]. Os benefícios desta interação são evidentes tais como estimular a motivação e trocas de conhecimento/experiência e auxiliar na discussão de problemas e dificuldades comuns dos estudantes, contribuindo para um ambiente não só agradável para a aprendizagem, mas principalmente tornando-a efetiva.

Uma das principais dificuldades em promover tal interação é definir o agrupamento (formação de grupos) de estudantes mais adequado, visando a realização de tarefas conjuntas, a fim de que alguma metodologia de ensino seja atendida e progressos constatados. Em ambientes presenciais, esta formação de grupos ocorre de maneira natural sem interferência do professor/tutor, já que os alunos se agrupam conforme conhecimento prévio e afinidades. Entretanto, em ambientes virtuais, este conhecimento e afinidades não estão disponíveis, sendo assim, deve ser fornecido algum suporte para identificação de algumas características que permitam aos estudantes e/ou ambiente definir grupos.

O modelo Big Five é uma organização hierárquica de traços de personalidade em cinco dimensões: abertura (O), conscienciosidade (C), extroversão (E), afabilidade (A) e neuroticismo (N) [De Raad 2000]. O modelo tem sido largamente utilizado na formação de grupos, pois cada dimensão agrupa elementos de personalidade compatíveis com as relações sociais e, portanto, com o agrupamento de pessoas. Estes traços podem ser identificados através de textos em linguagem natural, conforme pode ser observado em [Pennebaker et al. 2001], [Paim et al. 2016] e [Machado et al. 2015]. Desta maneira, é possível evitar o uso de inventários.

Entretanto, observa-se que embora o processamento de linguagem natural (PLN), atualmente, esteja altamente facilitado e fortemente explorado em diferentes aplicações, ainda não permite identificação de características psicológicas com alta acurácia. Assim, com o objetivo de constatar as principais causas desta dificuldade, duas equipes de pesquisadores observaram e compararam resultados sobre a identificação destes traços a partir de textos escritos. Os pesquisadores também conduziram estudos com o objetivo de aplicar os traços de personalidade detectados nos alunos na formação de grupos para apoiar a aprendizagem colaborativa. Assim, este trabalho apresenta e analisa os estudos realizados por estas equipes em relação à identificação dos traços e à formação de grupos.

Este artigo está organizado como se segue. A Seção 2 apresenta uma breve revisão da literatura relativa ao uso da teoria *Big Five* na formação de grupos para aprendizagem colaborativa, bem como reflexões sobre o uso de PLN para identificação dos traços. O referencial teórico sobre este estudo é apresentado na Seção 3. A Seção 4 descreve os experimentos conduzidos e compara os resultados obtidos, utilizando tanto bases de dados disponibilizadas para este fim como bases construídas pelos grupos de pesquisa. As conclusões e trabalhos futuros se encontram na Seção 5.

2. Trabalhos Relacionados

Uma das formas de conhecer características dos alunos para aplicar na formação de grupos é avaliar seus traços de personalidade e vários trabalhos apontam para sua influência no desempenho dos grupos. O estudo de [Akhtar et al. 2015], por exemplo, aponta conscienciosidade como agindo diretamente na motivação interna do indivíduo para atingir uma meta, alta afabilidade com maior chance de se envolver em trabalho de grupo do que aqueles com baixa afabilidade, e abertura, que segundo seus resultados, é o segundo preditor mais forte para o engajamento no trabalho.

Em Informática na Educação uma iniciativa utilizando Personalidade é a ferramenta *Group Recommender* [Nunes et al. 2010], que possibilita a criação de equipes de trabalho eficientes coordenadas por tutores, considerando as características do tutor e a similaridade de Traços de Personalidade dos alunos. Para utilização do sistema, os

usuários/alunos têm que responder ao *Personality Inventory* PV1.0 do qual se extrai o modelo de personalidade e, por meio deste, são gerados os dados necessários para a criação das equipes. O estudo de caso de [Rutherford 2006] aponta grupos heterogêneos quanto à personalidade como melhores. Em grupos homogêneos, todos terão os mesmos pontos fortes e as mesmas fraquezas, assim, ambos são multiplicados pelo número de integrantes. Em um grupo heterogêneo, haverá uma maior variedade de habilidade e fraquezas, podendo ajudar o grupo a gerenciar melhor os problemas que aparecem.

Ao avaliarem a capacidade de um indivíduo se conectar com novas pessoas ampliando sua rede de convívio social, [Roberts et al. 2008] apontam extroversão e neuroticismo como muito significativas para a socialização: a extroversão como um preditor da atividade social, como pedir ajuda e buscar novas conexões, e neuroticismo, o oposto. Já [Altanopoulou and Tselios 2015] sugerem a distribuição dos indivíduos com alta conscienciosidade entre os grupos, a fim de estimular a habilidade do grupo de cumprir o prazo proposto, e evitar formar grupo com muitos indivíduos com alto grau de extroversão, uma vez que tendem a se distrair mais com a interação social. De acordo com os resultados de [Bozionelos 2017], apesar de se esperar que indivíduos com alto neuroticismo tivessem interações ruins, tanto indivíduos com alto quanto baixo graus de extroversão se mostraram bons para a troca de recursos. O mesmo foi observado para abertura e afaabilidade. Aqueles com mais desvantagem em relação à interação para troca de recursos foram aqueles que obtiveram escores intermediários.

O estudo conduzido por [Carro and Sanchez-Horreo 2017] verificou a influência da Personalidade e dos Estilos de Aprendizagem na Educação Colaborativa, em um estudo de caso realizado com estudantes de Engenharia de Computação. Os dados dos alunos foram coletados por meio de atividades realizadas em ambiente virtual e com a aplicação do questionário NEO-FFI em conjunto com a identificação de inteligência por meio do teste *Primary Mental Abilities* (PMA) e dos estilos de aprendizagem com o modelo Felder-Silverman. Como resultado do trabalho, os autores indicam que a identificação do perfil de personalidade pode indicar os estudantes ou grupos que podem falhar nas atividades, bem como colaborar na construção de sistemas adaptativos com a formação de grupos dinâmicos.

Em um ambiente virtual de aprendizagem (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. Machado et al. [Machado et al. 2015] 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 (idioma Português do Brasil) e encontrou algumas relacionadas aos traços. [Paim et al. 2016] desenvolveram 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 estimar a personalidade por meio de textos escritos em Português. Em relação a modelos de detecção automática de características de personalidade a partir do texto, foi verificada a predominância da utilização de aprendizado de máquinas, como nas propostas de [Paim et al. 2016], [Tandera et al. 2017] e [Majumder et al. 2017].

3. Formação de Grupos e Personalidade

Aprendizagem colaborativa envolve a construção conjunta de significado por meio da interação com outros indivíduos e pode ser caracterizada por uma união para atingir uma meta compartilhada. A colaboração efetiva, então, envolve processos sociais por meio dos quais um grupo de alunos trabalham juntos para completar uma tarefa [Alavi 1994], enquanto é exposto a diferentes pontos de vista, questionando sua compreensão inicial e motivando sua aprendizagem. Apoiar os alunos em tarefas colaborativas ajuda, por exemplo, no desenvolvimento de pensamento crítico, auto-reflexão e responsabilidade por sua própria aquisição de conhecimento.

Apesar dos benefícios que podem ser atingidos, o processo de colaboração precisa de suporte para ser efetivado. Dentre os vários fatores que podem influenciar o processo de colaboração, a composição do grupo é considerada crucial para desencadear interações produtivas entre os membros do grupo [Magnisalis et al. 2011]. Como as características dos membros do grupo influenciam na maneira como trabalham juntos para alcançar uma meta, a formação do grupo baseada nas características individuais do aluno, como a personalidade, é essencial para permitir a aquisição de conhecimento através de atividades colaborativas [Manske et al. 2015].

A personalidade pode ser descrita como um conjunto de diferenças individuais que são afetadas pelo desenvolvimento de um indivíduo: valores, atitudes, memórias pessoais, laços sociais, hábitos e habilidades [McAdams and Olson 2010, Michel et al. 2004]. Uma das principais abordagens utilizadas para o estudo da personalidade humana é a teoria dos traços de [Kassin 2003], reconhecida como uma das maiores realizações da Psicologia [Deary 2009]. Com posições próprias, diversos autores indicaram preferências por modelos que contêm um diferente número de dimensões, sendo que o modelo mais aceito e difundido nesta área é o *Big Five Personality Traits*, que tem sido amplamente utilizado em estudos da personalidade humana e na identificação dos indivíduos em função dos seus traços de personalidade. Foi originalmente concebido por [Galton 1949], tendo suas raízes nas hipóteses léxicas para identificação de traços de personalidade, sendo baseado nas cinco dimensões: *Openness to Experience* (O); *Conscientiousness* (C); *Extraversion* (E); *Agreeableness* (A) e *Neuroticism* (N).

A utilização de questionários investigativos e questionários de autoavaliação, também conhecidos como inventários, tem dominado o campo de avaliação da personalidade. Conforme trabalho de renomados pesquisadores [Meyer et al. 2001] da área, as medidas padronizadas e referenciadas por normas (conjunto de questões, métodos de administração, pontuação), ou seja, questionários de avaliação, são os mais válidos e confiáveis métodos atualmente disponíveis para avaliar os construtos de personalidade. O *Personality Inventory- revised- NEO-PI-R* [Costa and McCrae 1992] pode ser considerado um dos mais relevantes trabalhos na área de identificação manual de perfil de personalidade, pois não somente demonstra grandes propriedades psicométricas, como também consegue acomodar construtores já endereçados pelas métricas existentes de traços de personalidade. Este questionário é baseado em respostas (cinco pontos da escala *Likert*) para cada uma das questões.

Alternativamente ao uso dos questionários, iniciativas de identificação automatizada têm sido apresentadas. Neste cenário, é verificada a necessidade de técnicas para estruturação do texto, que por natureza é uma fonte de informação não estruturada, de

maneira a viabilizar a aplicação de processos computacionais de identificação. Léxicos têm sido empregados para a obtenção de características do texto, através da associação de palavras a certas categorias pré-definidas, que podem incluir as emoções, no caso dos léxicos afetivos. LIWC (*Linguistic Inquiry and Word Count*) é um software que permite o cálculo da frequência de ocorrência de diferentes categorias de palavras em um conjunto de textos [Pennebaker et al. 2001]. Esta ferramenta baseia-se no dicionário LIWC, que é um léxico contendo cerca de 4.500 palavras previamente mapeadas em uma ou mais categorias diferentes, dentre as dezenas de categorias existentes neste dicionário. Além do dicionário no idioma Inglês, a versão do LIWC2007 incorporou uma versão do dicionário em Português [Balage Filho et al. 2013].

Um outro método utilizado para estruturação de texto é a associação da frequência de ocorrência das palavras encontradas no texto com vetores numéricos. O modelo de predição *n-gram* determina a frequência de ocorrência dos *n* elementos do texto, gerando as denominações *unigram*, *bigram* e *trigram*, quando estes elementos são associados a uma vizinhança de um, dois ou três elementos, respectivamente. Na área de classificação automática de documentos, a utilização de informações estatísticas da ocorrência de *n-gram* produz relevantes ganhos neste processo [Cavnar et al. 1994]. Para a obtenção da frequência de ocorrência de cada *n-gram* podem ser utilizadas técnicas como a *term frequency-inverse document frequency* (TF-IDF), uma medida estatística que tem como objetivo indicar a relevância de um termo em um conjunto de textos.

4. Experimentos

Nesta seção são apresentados dois experimentos. O primeiro com o objetivo de detectar automaticamente os traços de personalidade de alunos, para empregar na formação de grupos de colaboração. O segundo com o objetivo de avaliar a influência do perfil dos alunos no desempenho do grupo, avaliado através das interações e notas obtidas em atividade colaborativa. Os perfis dos alunos foram identificados através da aplicação do inventário de personalidade de 44 questões em sua versão traduzida para o Português [Andrade 2008] e com validade testada no Brasil.

4.1. Experimento de Detecção de Personalidade

Este experimento foi realizado junto a alunos de ensino superior, na área de Educação, em uma universidade pública no estado do Paraná. Os voluntários, alunos regulares em uma disciplina de ensino presencial, com apoio do ambiente *Moodle*, foram incentivados a utilizar a atividade de Fórum de Discussão. Ao conjunto de 45 alunos, que concordaram em participar do experimento, foi aplicado o questionário e, a partir das respostas, foi obtida a identificação dos traços de personalidade, no modelo *Big Five*, para ser utilizada como referência no experimento. As informações presentes no ambiente *Moodle*, correspondentes às mensagens inseridas pelos alunos, foram utilizadas para a identificação automatizada do traço de personalidade a partir das palavras presentes no texto, utilizando o modelo apresentado por [Buiar et al. 2017]. Com a utilização da ferramenta LIWC2015, foi obtida a representação dos textos nas 80 categorias léxicas oferecidas pelos dicionários utilizados pela ferramenta. Com isto, o texto foi estruturado, permitindo a aplicação dos processos de classificação. Os resultados obtidos são apresentados na Tabela 1, onde são indicados os classificadores que obtiveram a melhor acurácia para cada dimensão de

personalidade, bem como os valores de Acurácia, Erro, Sensibilidade (Sen), Especificidade (Esp), Valor Previsto S (VPS), Valor Previsto N (VPN) e Desvio Padrão do Erro (DPErro). O critério *kfolds* de separação da base em treinamento e teste foi utilizado, com $k=3$. Estes resultados indicam melhores resultados no caso da dimensão *Openness* e

Tabela 1. Resultados do Experimento.

Dim	Acurácia	Erro	Sens	Esp	VPS	VPN	DPErro	Classificador
O	89%	11%	97%	16%	90%	16%	0,02	kNN
C	57%	43%	68%	42%	62%	50%	0,03	AdaBoost
E	66%	34%	72%	60%	70%	62%	0,09	Random Forest
A	71%	29%	50%	85%	72%	72%	0,08	Gaussian NB
N	61%	39%	92%	23%	61%	43%	0,09	ML Perceptron

um maior desvio padrão do erro nas três últimas dimensões. Este experimento demonstrou a obtenção de resultados coerentes com os valores encontrados na literatura, para processos de identificação perfil de personalidade, a partir do texto, mas em função do pequeno número de registros na base original, foi verificada a necessidade de realização de um experimento com uma base maior de alunos.

4.2. Experimento de Formação de Grupos

Os dados utilizados foram coletados de estudantes (nível médio ou superior), que aceitaram participar da pesquisa, por meio de respostas ao questionário e cinco perguntas adicionais que forneceram texto para o treinamento do modelo de predição dos traços de personalidade. A coleta das respostas foi feita com o uso de uma escala *Likert* de 5 pontos. As perguntas adicionais eram cenários nos quais o aluno deveria descrever sua possível resposta ou reação em cada um deles. 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.

A resposta ao questionário foi facultativa e um total de 236 alunos, com idade entre 15 e 56 anos, responderam ao questionário, sendo 106 do sexo feminino e 130 do sexo masculino. Destes, 35 alunos de ensino médio técnico em Informática, 173 de ensino superior e 28 alunos de pós-graduação. O número de alunos com ensino superior ou pós-graduação considerou tanto os que já haviam concluído quanto aqueles que ainda estavam cursando. Os participantes eram alunos da Faculdade de Computação e da Faculdade de Gestão e Negócios da Universidade Federal de Uberlândia (UFU), da Faculdade Pitágoras de Uberlândia, do Centro Universitário do Triângulo (Unitri) e do Instituto Federal Goiano - Câmpus Campos Belos. Todos de cursos relacionados à tecnologia da informação, administração e gestão da informação.

Os indicadores analisados, similarmente ao experimento anterior, são apresentados na Tabela 2 para este segundo experimento. Observa-se que quanto maior o valor de VPN (ou VPS), melhor a capacidade do classificador acertar quando prevê N (ou S). Sen e Esp medem a porcentagem da amostra de S e N, respectivamente, que o classificador foi capaz de estimar corretamente. Como a base foi construída por mais amostras de S do que N (mais que 80% de S para *Openness*, *Agreeableness* e *Conscientiousness*), os classificadores que obtiveram maior acurácia foram aqueles onde $Esp=0$ e $VPN=0$. Por

isso, além de considerar erro e acurácia, os algoritmos da Tabela 2 foram selecionados considerando os valores de Sen, Esp, VPS e VPN.

Tabela 2. Resultados do Experimento.

Dim	Acurácia	Erro	Sen	Esp	VPS	VPN	DPErro	Classificador
O	83%	17%	95%	8%	86%	25%	0.04	IBK
C	85%	15%	98%	9%	87%	42%	0.03	RandomForest
E	63%	37%	74%	40%	71%	45%	0.03	Perceptron
A	85%	15%	97%	12%	87%	50%	0.03	AdaBoost
N	57%	43%	51%	62%	55%	59%	0.06	J48

A Tabela 3 resume os resultados obtidos de 13 grupos trabalhando colaborativamente. A performance dos grupos foi observada em relação à completa tarefa (CT), interação (I) via conversa, nota atribuída pelo professor (Nota) e todos contribuem para a solução (TCS) da tarefa proposta. Tam é o número de alunos no grupo. O perfil do aluno é definido pela cadeia de valores na seguinte ordem *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness* e *Neuroticism*. Os valores de cada traço são representados por um caractere que indica uma das faixas para o escore obtido com o questionário. A faixa 1, representada por o, c, e, a e n, indica escore menor que 38,5%, a faixa 2, representada por -, significa escore entre 38,5% e 62,5% inclusive, e a faixa 3 (O, C, E, A e N) representa escores maiores que 62,5%. Por exemplo, o aluno com perfil OCE-N obteve escore médio (faixa 2) para *Agreeableness* e escores altos (faixa 3) para os demais.

Tabela 3. Performance grupos e perfil de alunos do grupo.

Grupo	CT	I	Nota	TCS	Tam	Aluno1	Aluno2	Aluno3
1	1	1.00	0.50	1	2	OCE - N	- cE - N	
2	1	1.00	0,66	1	3	- - - - N	- C- An	- CEAN
3	1	1.00	0,73	1	3	oCE - N	OCE - N	- CE - -
4	0	1.00	0,36	0	3	- CEA -	- C - - -	- - - A -
5	1	1.00	0,53	1	3	OCEAn	OCEA -	OCEA -
6	1	1.00	0,83	1	3	- - E - n	OC - - n	- C - - n
7	1	1.00	0.60	1	3	OC - - n	O - EA -	- - - An
8	1	1.00	0.50	1	3	- - E - N	oC - - n	- - E - N
9	1	0.50	1.00	1	3	O - - A -	Ocea -	OC - A -
10	1	0.60	0.59	0	3	OCEAn	- - e - N	OC - - -
11	1	0.53	0.64	0	3	- C- An	OC - An	OCE - n
12	1	0.67	0.69	0	3	- - - - N	OC - - N	- CEAn
13	1	0.67	0.65	1	3	- C - An	OC - - n	- Ce - N

Na avaliação dos grupos, como o objetivo era formar grupos bons para colaboração, foram considerados como bons grupos aqueles em que todos os membros trabalharam na solução da tarefa proposta e, além disso, havia comunicação que possibilitasse a interação social. Os grupos formados por membros que tinham valores altos ou baixos de conscienciosidade, afabilidade e neuroticismo pareciam trabalhar bem juntos. O mesmo ocorreu quando a maioria do grupo tinha alta conscienciosidade, abertura e extroversão, bem como naqueles formados por indivíduos que apresentavam baixa ou

média extroversão. Por outro lado, nos grupos onde a maioria tinha escore médio ou médio e alto, alguns foram bons outros não, o que poderia indicar que a performance foi influenciada por outras características dos membros do grupo.

Em geral, nos grupos em que os membros demonstraram maior participação e alta interação, a maioria apresentava conscienciosidade alta. Por outro lado, os 4 grupos em que nem todos os alunos participaram da solução (TC=0) também foram formados por membros conscienciosos. Nestes grupos, a falta de colaboração pode ser explicada pela combinação dos demais traços. Nos grupos 10 e 12, por exemplo, há membros altamente extrovertidos e com baixo neuroticismo, que tendem a apresentar excesso de entusiasmo, assertividade e autoconfiança, o que pode intimidar aqueles com alto neuroticismo ou baixa extroversão. Os grupos onde a maioria dos membros apresentou escore médio na maioria das dimensões, como no grupo 4, podem não funcionar bem.

5. Conclusão

Este artigo descreveu a aplicação do modelo *Big Five* na formação de grupos de alunos, utilizando textos em linguagem natural na identificação dos perfis de personalidade. A identificação de perfil, a partir do texto, apresenta um grande desafio computacional. As técnicas mais utilizadas estão baseadas no aprendizado de máquina, e uma série de estratégias de representação do texto de forma estruturada, a fim de permitir a realização de processos de classificação. Mesmo com grandes esforços na utilização de classificadores mais complexos e técnicas de representação mais aprimoradas, a melhoria nos indicadores de classificação, em especial a acurácia, não avançam na mesma proporção.

A dificuldade da obtenção de bases de treinamento para os classificadores, os ruídos presentes no processo de identificação a partir de questionários de avaliação e a quantidade de palavras presentes nos textos utilizados interferem de forma direta nos resultados. Nos experimentos realizados, o uso de *n-gram* durante o processamento da linguagem natural foi uma alternativa explorada, visando melhorar a identificação das categorias do dicionário. Entretanto, os resultados não justificaram a complexidade deste acréscimo.

Em um experimento inicial foi verificada a aplicação de um processo de identificação em ambiente educacional, mesmo com uma quantidade reduzida de alunos que participaram do experimento. A partir desta referência foi conduzido um segundo experimento em maior escala, onde a identificação da personalidade foi utilizada para a verificação da formação de grupos. Os resultados demonstraram que uma metodologia de identificação da personalidade dos alunos pode oferecer suporte relevante para tutores e estudantes em AVAs. Contudo, considerando-se que a quantidade e a diversidade de texto são de suma importância para a classificação dos traços e que as características qualitativas dos grupos formados estão em conformidade com a literatura, os resultados foram promissores.

Assim, como perspectiva futura, os pesquisadores destes grupos acreditam que melhoria significativa está associada ao uso de outras métricas que, em conjunto com o modelo *Big Five*, permitam a identificação de outros elementos relevantes para formação de grupos como por exemplo, o desempenho do grupo, número de interações entre os membros e o grau de participação no desenvolvimento da atividade colaborativa proposta. Os resultados deste trabalho também abrem espaço para o desenvolvimento de uma fer-

ramenta para formação de grupos, usando o perfil do aluno detectado automaticamente a partir de textos. Tal ferramenta deve ser desenvolvida como trabalho futuro.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Referências

- Akhtar, R., Boustani, L., Tsivrikos, D., and 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. and 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., Pardo, T. A. S., and Aluísio, S. M. (2013). An Evaluation of the Brazilian portuguese LIWC Dictionary for Sentiment Analysis. *9th 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.
- Buiar, J., Pimentel, A., and Oliveira, L. (2017). Identificação de estilo de aprendizagem: Um modelo de inferência automatizado baseado no perfil de personalidade identificado nos textos produzidos pelo aluno. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 28, page 1157.
- Carro, R. M. and Sanchez-Horro, V. (2017). The effect of personality and learning styles on individual and collaborative learning: Obtaining criteria for adaptation. In *Global Engineering Education Conference (EDUCON), 2017 IEEE*, pages 1585–1590. IEEE.
- Cavnar, W. B., Trenkle, J. M., et al. (1994). N-gram-based text categorization. *Ann arbor mi*, 48113(2):161–175.
- Costa, P. T. and McCrae, R. R. (1992). Neo personality inventory–revised (neo-pi-r) and neo five-factor inventory (neo-ffi) professional manual. *Odessa, FL: Psychological Assessment Resources*.
- De Raad, B. (2000). *The Big Five Personality Factors: The psycholexical approach to personality*. Hogrefe & Huber Publishers.
- Deary, I. J. (2009). The trait approach to personality. *The Cambridge handbook of personality psychology*, pages 89–109.
- Everson, M. (2011). What can you gain by teaching online? *eLearn Magazine*, 2011(1):3.
- Galton, F. (1949). *The Measurement of Character*. Prentice-Hall, Inc.
- Kassin, S. M. (2003). *Essentials of psychology*. Prentice Hall.

- Kop, R. (2011). The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *The International Review of Research in Open and Distance Learning*, 12(3):19–38.
- Machado, A. A., Longhi, M. T., Nunes, M. A. S. N., and Pardo, T. A. S. (2015). Personalitatem 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., and Karakostas, A. (2011). Adaptive and intelligent systems for collaboration learning support: A review of the field. *IEEE Transactions on Learning Technologies*, 4:5–20.
- Majumder, N., Poria, S., Gelbukh, A., and Cambria, E. (2017). Deep learning-based document modeling for personality detection from text. *IEEE Intelligent Systems*, 32(2):74–79.
- Manske, S., Hecking, T., Chounta, I. A., and Hoppe, H. U. (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.
- McAdams, D. P. and Olson, B. D. (2010). Personality development: Continuity and change over the life course. *Annual review of psychology*, 61:517–542.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., and Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American psychologist*, 56(2):128.
- Michel, W., Shoda, Y., and Smith, R. (2004). Introduction to personality: Toward an integration.
- Nunes, M., Moraes, D., and Reinert, D. (2010). Personality inventory-pv 1.0 (portuguese version). *Instituto Nacional de Propriedade Industrial.(software register INPI-12093-0)*.
- Paim, A., Camati, R., and 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.
- Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001.
- Roberts, S. G. B., Wilson, R., Fedurek, P., and Dunbar, R. I. M. (2008). Individual differences and personal social network size and structure. *Personality and Individual Differences*, 44:954–964.
- Rutherford, R. H. (2006). Using personality inventories to form teams for class projects – a case study. In *Proceedings of SIGITE '06 Proceedings of the 7th conference on Information Technology Education*, pages 9–14.
- Tandera, T., Suhartono, D., Wongso, R., Prasetio, Y. L., et al. (2017). Personality prediction system from facebook users. *Procedia Computer Science*, 116:604–611.