

Melhorar a previsão inicial de insucesso escolar utilizando análise de sentimentos na auto - comentários avaliadas

LC Yu ¹ | CW Lee ¹ | HI Pan ¹ | CY Chou ² | PY Chao ³ | ZH Chen ⁴ | SF Tseng ¹ |

CL Chan ¹ | KR Lai ²

¹ Departamento de Gestão da Informação, Yuan Zé

Universidade, Taoyuan City, Taiwan

² Departamento de Ciência da Computação e Engenharia, Yuan Zé

Universidade, Taoyuan City, Taiwan

³ Departamento de Comunicação da Informação, Yuan Zé

Universidade, Taoyuan City, Taiwan

⁴ Instituto Universitário de Informação e Educação em

Computação, National Taiwan Normal University, Taipei City, Taiwan

Correspondência

Liang - Chih Yu, do Departamento de Gestão da Informação, Yuan

Zé Universidade, n° 135, Yuandong Road, Zhongli District,

Taoyuan City

320, Taiwan.

Email: lcyu@saturn.yzu.edu.tw

informações sobre financiamento

Ministério da Ciência e Tecnologia, Grant / Números de

adjudicação: MAIS 105 - 2218 - E - 006 - 028 e 105 MAIS - 2221 - E - 155 - 059

- MY2

Abstrato

Este estudo apresenta um modelo para a identificação precoce de estudantes que estão propensos a falhar em um curso acadêmico.

Para melhorar a precisão da previsão, análise de sentimentos é usado para identificar informações afetiva de texto - auto base - comentários

avaliadas escritos por alunos. Os resultados experimentais demonstraram que a adição de informações sentimento extraído de auto **estudante - avaliações produz uma melhoria significativa no início de - qualidade predição fase. Os resultados também indicam o início limitado - encenar valor preditivo de dados estruturados, como a realização trabalhos de casa, a frequência, e os tipos de exames,** devido à escassez de dados, no início do curso. Assim, a aplicação de análise de sentimentos de dados não estruturados (por **exemplo, auto - comentários de avaliação) podem desempenhar um papel importante na melhoria da precisão do início - previsões de palco. As descobertas atuais educadores com uma oportunidade de proporcionar aos alunos reais - feedback em tempo e apoio para ajudar os alunos a se tornar auto - alunos regulamentadas. Usando os resultados de exploração por melhorias no ensino e iniciativas de aprendizagem é importante para manter o desempenho dos alunos e da eficácia do processo de aprendizagem.**

PALAVRAS-CHAVE

previsão inicial, análise de aprendizagem, análise de sentimentos, dados não estruturados

1 | INTRODUÇÃO

estimativa de desempenho acadêmico é importante não só para ajudar os estudantes a assumir o **controle de sua própria aprendizagem e tornar-se auto - regulada alunos, mas também para permitir** que os educadores para identificar estudantes em risco e intervir mais cedo para reduzir o risco de fracasso. avaliação formativa e sumativa são geralmente adotado para estimar o desempenho acadêmico. Os dados utilizados para tais avaliações contêm informações valiosas, tais como tendências e padrões relacionados com o processo de aprendizagem nas práticas educativas, que podem ser usados para compreender o estado de aprendizagem. No entanto, essas avaliações são **complexos eo tempo - demorado de implementar, e detalhes importantes podem ser negligenciados.** Uma abordagem alternativa é a aplicação de técnicas computacionais automáticas para analisar os dados do aluno e para facilitar e melhorar o processo de avaliação.

Com o surgimento de técnicas e avanços na computação grandes de análise de dados, análise de aprendizagem emergiu como um foco de investigação crescente no domínio da **educação. Aprender análise procura identificar aprendizagem - padrões relacionados em contextos** educativos e, assim, melhorar

compreensão dos processos educacionais, avaliar os resultados da aprendizagem, e fazer previsões de desempenho (Jayaprakash, Moody, Lauria, & Baron, 2014). Muitos de análise de aprendizagem algoritmos e métodos úteis foram desenvolvidos, incluindo clustering, classificação, árvores de decisão, a mineração de regras de associação, regressão e visualização, e estas abordagens têm sido amplamente aplicado no desenvolvimento de modelos para prever o insucesso escolar. Bayer, Bydzovská, Géryk, Obsvac e Popelinsky (2012) focada em prever o abandono e insucesso escolar, utilizando dados baseados em comportamentos **sociais dos alunos, como estudante - para - aluno e aluno - para - relações de professores e interações** identificadas por análise de redes sociais. Tais comportamentos foram então processadas por vários métodos de aprendizado de máquina (por exemplo, SVM, Naive Bayes, e J48 decisão árvore) para melhorar a precisão da previsão. Shahiri, Husain, e Rashid (2015) forneceu uma visão geral sobre a previsão graus acadêmicos finais utilizando várias técnicas analíticas. Smith, Lange, e Huston (2012) desenvolveu modelos preditivos utilizando ingênua classificação Bayes **para identificar a - estudantes de risco com base em dezenas de variáveis-chave. Estes estudos** demonstraram capacidade preditiva eficaz

e identificou variáveis-chave para prever com precisão o fracasso. No entanto, estes estudos não estão preocupados com dois aspectos críticos: (a) a importância do momento e (b) a relação de aluno estado emocional e sucesso acadêmico.

Claramente, a intervenção precoce é um fator chave na prevenção de insucesso escolar em **pelo - estudantes de risco** (Jayaprakash et al., 2014), e Santana, Costa, Fonseca, Rego, e de Araújo (2017) salientou a importância de prever a falha com antecedência suficiente para permitir intervenções pedagógicas para prevenir o insucesso escolar. Para conseguir a identificação precoce **de pelo - estudantes de risco**, Santana et al. (2017) tentaram melhorar a qualidade dos dados usando uma sequência de passos de pré-processamento de dados, tais como a extração de dados, de limpeza e de transformação, e melhorar a eficácia modelo utilizando um algoritmo para seleccionar atributos que melhor se adaptam ao modelo. No entanto, Pritchard e Wilson (2003) observou que o sucesso acadêmico tem sido tradicionalmente prevista usando dados estruturados, como variáveis demográficas e acadêmicos, mas estas variáveis claramente não são responsáveis por toda a variação no sucesso acadêmico. Para ir além preditores tradicionais, valiosas fontes de dados não **estruturados precisam ser incluídos para conseguir a identificação precoce de pelo - estudantes de risco**.

Auto - avaliações escritas por estudantes são uma fonte de dados não estruturados útil, contendo informação emocional rica que pode ajudar a iluminar os estados emocionais dos alunos. Sorour, Goda, e Mine (2015) enfatizou o potencial de comentários de estudantes para refletir as atitudes dos alunos de aprendizagem, compreensão assunto e dificuldade lição, ajudando professores ajustar e aperfeiçoar as atividades de aprendizagem dos alunos. Pekrun, Goetz, Titz, e Perry (2002) sugeriu que as emoções dos alunos, enquanto frequentar as aulas, **em wileyonlinelibrary.com | VU** são variáveis importantes para determinar o sucesso acadêmico. Vários estudos têm confirmado o impacto dos estados emocionais no desempenho acadêmico (D'Mello et al, 2008; Luo, Sorour, Goda, e Mine, 2015; Sorour, Mina, Goda, e Hirokawa, 2015).

Este estudo apresenta um modelo para a identificação precoce de estudantes que estão propensos a falhar em um curso acadêmico. Para melhorar a precisão da previsão inicial, análise **de sentimentos é usado para identificar informações afetiva dentro do texto - auto base - avaliações** escritas pelos alunos. Esta informação afectiva é então combinado com dados estruturados (por exemplo, a frequência, a realização trabalhos de casa, e em - participação da classe) como entrada **auto - comentário avaliado figura colonda pode ser visto** para construir um modelo de previsão. algoritmos de aprendizado de máquina são usados para determinar se o modelo pode melhorar o desempenho de previsão para além usando dados estruturados sozinho. Um total de 181 estudantes universitários participaram em experimentos onde eles foram convidados

para produzir auto escrito - avaliações após cada aula, produzindo 2.476 comentários. Cada comentário foi anotado manualmente com um resultado discreto chamado de intensidade **sentimento ou força entre 1 e 9, com base na auto - escala de avaliação Assessment Manikin** (Lang, 1980). Uma pontuação de 1 denota um comentário extremamente negativo, enquanto que 9 denota um comentário extremamente positivo, e 5 denota um sentimento neutro. Como mostrado na Figura 1, um gráfico de linhas é usado para representar as emoções do aluno ao longo do tempo, **permitindo que os professores fornecem verdadeira - feedback em tempo para os alunos com** flutuante ou consistentemente baixas humores e emoções. Por exemplo, um comentário estudante mencionou que o estudante se sentiu confuso com novos termos excessivos utilizados durante a terceira semana de aula. Este comentário é dada uma pontuação de intensidade emocional de 4, o que indica que é um pouco negativa. Em reação a isso, o professor pode pedir a um assistente de ensino para fornecer suplementar explicação antes da próxima aula.

Para destacar informações úteis e tomada de decisão de apoio, nós desenvolvemos uma **auto - sistema regulado chamada a dinâmica de auto diagnóstico e - sistema regulado (DDS)**, utilizando técnicas gráficas para conseguir a visualização da informação. Em contextos educativos, representações visuais de atividades do curso e registros de uso podem proporcionar aos educadores e estudantes com uma visão geral do status de aprendizagem (Romero & Ventura, 2010). O sistema DDS contém dados estruturados (por exemplo, a frequência, a realização trabalhos de **casa, e em - participação da classe**) e **dados não estruturados (self - Comentários avaliados) coletados de plataforma de aprendizagem virtual proprietária da universidade, que podem fornecer verdadeira - sugestões** de tempo de acordo com o estado de aprendizagem do aluno para promover a autonomia do aluno.

O restante deste trabalho está estruturado da seguinte forma. A Seção 2 apresenta trabalhos anteriores relacionados com a nossa abordagem proposta. A Seção 3 apresenta fontes de dados e métodos de pesquisa utilizados neste estudo. Secção 3.2.1 resume os resultados **experimentais e introduz o auto proposto - sistema regulado. Resultados e conclusões estão** finalmente apresentado na Seção 4.1.

2 | REVISÃO DA LITERATURA

2.1 | emoções dos alunos e aprendizagem

emoções dos alunos pode afetar profundamente seus pensamentos, motivações e ações. D'Mello et al. (2008) enfatizou que a detecção de emoções aluno é importante na promoção da aprendizagem eficaz devido aos vínculos indissolúveis entre cognição e emoção. As emoções positivas podem aumentar o interesse dos alunos e motivação na aprendizagem. Pekrun et al. (2002) definiu as emoções do aluno como o emocional

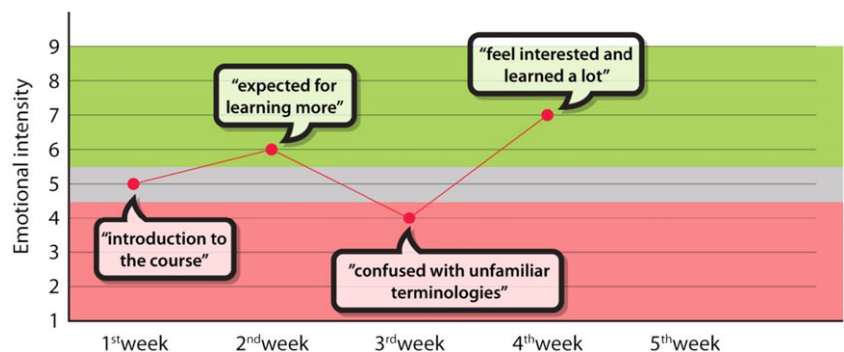


FIGURA 1 Estudante intensidade emocional extraído de

estudantes de experiência têm em ambientes acadêmicos quando frequentar as aulas, estudar e fazer provas. estados emocionais dos alunos estão significativamente relacionadas com a sua motivação e desempenho acadêmico. Por exemplo, a ansiedade do teste está negativamente relacionado ao aprendizado e realização. Altos níveis de stem test ansiedade de comparar auto - desempenho com Yu et al. (2016), a dimensão excitação é mais difícil de prever do que a dimensão de valência tanto para textos de língua chinesa Inglês e. No entanto, é ainda útil porque um - análise de valência dimensional centra-se na previsão de uma verdadeira - pontuação valorizado, chamado intensidade sentimento ou força (Kiritchenko, Mohammad, e Salameh, 2016), que pode fornecer ambos os sentidos positivos e negativos e níveis de intensidade para uso em análise emocional (Mishne & De Rijke, 2006; Nguyen, Phung, Dao, Venkatesh, e Berk, 2014).

2.2 | Análise de sentimentos

Os resultados da aprendizagem, motivações e atitudes são tipicamente avaliada utilizando exames e observando o desempenho do aluno em sala de aula. No entanto, ambas as abordagens podem ser tempo e trabalho intensivo, aumentando a necessidade de uma forma mais eficiente para fornecer avaliações contínuas e longitudinais de aprender eficácia e atitudes. Luo et al. (2015) observou que os comentários escritos por alunos após cada aula reflete suas atitudes de aprendizagem para com essa lição, compreensão do conteúdo do curso, e dificuldade percebida de se inclinar. Eles pediram aos alunos para escrever comentários após cada lição e, em seguida, extraído informações a partir desses comentários para rastrear continuamente o status aprendizagem dos alunos. técnicas de mineração de texto automatizados pode ser efetivamente usado para extrair e analisar grandes quantidades de dados textuais. Vários estudos (Calvo & D'Mello, 2010; Feldman, 2013; Yu et al., 2016) têm utilizado a análise sentimento de identificar automaticamente as informações afetiva de textos, que podem ser usados para reconhecer automaticamente estudantes estados emocionais. Na análise de sentimentos, estados afetivos são geralmente representados usando abordagens quer categóricas ou dimensionais.

A abordagem categórica representa estados afetivos como várias classes discretas tais como positiva, neutra, negativa e de Ekman (1992) seis emoções básicas (por exemplo, raiva, felicidade, medo, tristeza, desgosto e surpresa). Na base desta representação, várias aplicações práticas têm sido desenvolvidos, tais como aspecto - análise de sentimento baseado (Pontiki et al, 2014; Schouten & Frasincar, 2016), análise de sentimentos Twitter (Rosenthal et al, 2015; Saif, Fernandez, Ele, & Alani,

2013), detecção de spam opinião enganosa (Li, Ott, Cardie, & Hovy, 2014), e portabilidade crosslingual (Banea, Mihalcea, & Wiebe, 2013; Xu, Gui, Xu, Lu, e Wong, 2015).

Em contraste com a abordagem categórica, a abordagem dimensional representa afectar estados usando um valor numérico contínuo para várias dimensões. As dimensões mais vulgarmente utilizados incluem a valência - espaço excitação, a dois - abordagem dimensional, onde a dimensão de valência reflete o grau de sentimento positivo e negativo, ea dimensão excitação reflete o grau de calma e emoção. Qualquer expressão de sentimento pode ser representado como um

ponto em que a valência - excitação plano de coordenadas, reconhecendo a sua valência - classificações de excitação. Os aplicativos podem usar tais representações para fornecer mais fina - granulação (por exemplo, bens - valorizados) análise de sentimento (Kiritchenko & Mohammad, 2016). De

2.3 | visualização de informação

Simon (1996) observou que a solução de um problema significa simplesmente representando-o de modo a tornar a solução transparente. visualização de informação permite aos usuários avaliar e compreender os dados em um relance. Este ramo da computação gráfica e design de interface de usuário está preocupado com a apresentação eficaz de digital interativa ou animado

imagens

(Spence, 2001). Estas técnicas facilitar a análise de grandes quantidades de informação através da geração de uma representação visual dos dados. visualização de informação dos resultados da análise oferece uma visão geral dos resultados de aprendizagem de um aluno e suporta tomada de decisão eficaz por meio destacando informações úteis. Romero e Ventura (2010) enfatizou o benefício de visualização de informação de dados analisados em ambientes educacionais, especialmente para resolver tempo - demorado processo de interpretação.

Learner dados foram recolhidos e analisados para a produção de visualização de informações. Professores e alunos foram fornecidos com gráficos estatísticos resumindo presença dos alunos, a conclusão atribuição, o acesso aos recursos, resultados de exames, e estado emocional de oferecer uma visão geral do status de aprendizagem.

3 | MÉTODO DE PESQUISA

3.1 | Fonte de dados

Os dados foram obtidos a partir de 181 estudantes de graduação em dois Introdução separado para Ciência da Computação cursos durante o 19 - sessão de verão semana no Yuan Zé Universidade de Taiwan. Os alunos foram avaliados em termos de atendimento, conclusão lição de casa, em - participação da classe e exames no 5º, 11º e 18ª semana do curso. Uma classe de passagem para o curso é de 60. As pontuações para as duas classes estavam (média, o desvio padrão) = (57.2, 15.9) de classe A e (51.8, 15.3) de classe

B. Resultados da aprendizagem e distribuições de grau são apresentados na Tabela 1. Após cada aula, os alunos foram convidados a escrever auto -

avaliações para a reflexão sobre o seu estado inclinada pessoal. Um total de 2.476 tais observações foram recolhidos. Uma vez duplicados ou textos incompreensíveis tinham sido eliminados, a amostra totalizou 2.449 textos com um total de 126,318 caracteres chineses (sem pontuação), para uma média de 51,58 caracteres por texto. Um total de 12 observações continha menos de cinco caracteres, incluindo " 好! "(Boa!); " 有難度 "(É difícil de entender); " 老師真趣

"(O professor tem um bom senso de humor); e " 學習到很多 "(Aprendi muito nesta lição). apesar de sua

TABELA 1 Resultados da aprendizagem e distribuições de grau

	Num. de estudantes	Resultado de aprendizagem		pontuações	
		Num. de passagem (pontuação > 60)	Num. de falhar (pontuação < 60)	Pontuação média	Desvio padrão
Classe A	95	44	51	57,2	15,9
classe B	86	27	59	51,8	15,3
Total	181	71	110	54,6	15,8

brevidade, esses comentários continha informações úteis sobre o estado de aprendizagem do aluno. Na outra extremidade do espectro, com um total de 263 textos containedmore do que 100 caracteres. todos os auto - comentários de avaliação indicado o grau de compreensão e realização de cada aluno em cada etapa do curso.

A fonte de dados contém as seguintes informações de cada aluno: assiduidade, a conclusão lição de casa, em - participação da classe, e as emoções do aluno extraído do auto - comentários de avaliação. Os detalhes são apresentados na Tabela 2.

3.2 | Métodos de análise

3.2.1 | previsão emoção

análise de sentimentos é usado para extrair estudante emoção intensidade do auto - comentários avaliadas. Cada comentário foi avaliado por três anotadores voluntários 1-9 em termos de valência baseado no Auto - escala de avaliação Assessment Manikin (Lang, 1980). Os valores nominais foram então a média para produzir valência do comentário e categorizados a três tipos emoção diferente como (classificação de valência sob 4,5) negativo, neutro (classificação de valência entre 4,5 e 5,5) e positiva (classificação de valência acima de 5,5). A Tabela 3 mostra o número de observações e valência para vários tipos emoção de acordo com cada um anotador significa. Comparando os tipos de emoção avaliado por cada anotador mostra um kappa média de Cohen (Cohen, 1960) de 0,85 para confiabilidade inter.

No desenvolvimento de aplicações sentimento, léxicos afetivos e corpora com classificações valência são recursos úteis para a mineração opinião.

recursos afetivos são normalmente obtidos por qualquer auto - rotulagem ou anotação manual (Yu et al., 2016). Na auto - abordagem rotulagem, os usuários oferecer de forma proativa sentimentos e opiniões em resposta a navegação na web - conteúdo baseado. No método de anotação manual anotadores treinados são convidados a criar anotações afetivas para os recursos linguísticos específicos. Nós construímos um léxico afetivo Chinês (Yu et al., 2016) com classificações valência para cada palavra afetiva e é usado para extrair automaticamente emoções de auto - comentários de avaliação. A classificação valência de cada comentário foi calculado pela média das classificações de valência das palavras combinadas no léxico em que o comentário. Por exemplo, a Figura 1 mostra um comentário aluno da terceira semana no qual se constata,

" Estou confuso com terminologias desconhecidas. " A classificação de valência da palavra afetiva " confuso " no léxico é 4, o que significa aluno experimentou emoções negativas após esta lição. Em resposta à dificuldade citada com vocabulário novo, na Semana 4, o estudante foi fornecido com material suplementar e auto posterior do aluno - avaliação contou com o comentário: " Sinto-me interessado e aprendi muito. " A palavra afetiva " interessado " no léxico tem uma classificação de valência de 7, indicando o aluno experimentou emoções mais positivas depois de assistir a esta lição. Se duas ou mais palavras afetivas aparecem no auto - comentários avaliadas, a classificação valência do comentário dado é calculado pela média das classificações de valência de cada palavra. Uma vez que a classificação de valência para cada comentário é obtida, a classificação é convertido para um tipo de emoção classificada como negativa (classificação de valência sob

4,5), neutro (classificação de valência entre 4,5 e 5,5), ou positiva (classificação de valência superior a 5,5).

MESA 2 Fonte de dados

tipo de atributo		Nome do Atributo	Descrição
preditores	Os dados estruturados	grau Homework	Os graus de lição de casa em média
		Homework clique	Número de instrução lição de casa leitura
		Comparecimento	Número de classe assistir
		Discussão	Número de postdiscussion
	dados não estruturados de emoções dos alunos	valência positiva	Número de valência positiva
		valência negativa	Número de valência negativa
		valência neutra	Número de valência neutra
		Média de valência	Valence em média
Alvo		Performance acadêmica	Falha ou passado este curso

TABELA 3 Resultados de anotação

Valência		Annotator 1		Annotator 2		Annotator 3		Todos (média)	
tipo de emoção	gama Valence	Num. dos comentários (total: 2.449)	Mia de valência	Num. dos comentários (total: 2.449)	Mia de valência	Num. dos comentários (total: 2.449)	Mia de valência	Num. dos comentários (total: 2.449)	Mia de valência
Negativo <4.5		284	3.965	455	3.440	373	3.440	391	3.835
Neutro	4,5 - 5,5	1.596	4.995	1.162	5.000	1.331	5.000	1.316	5.037
Positivo > 5.5		569	6.037	832	6.380	745	6.600	742	6.145

O desempenho do léxico - abordagem à emoção previsão é avaliada usando precisão, definida como

$$\text{Precisão} = \frac{\text{número de comentários cujos tipos de emoção são corretamente identificados}}{\text{número total de comentários}} \quad (1)$$

3.2.2 | previsão de desempenho acadêmico

análise de dados abordagens para dados educacionais, máquina de vetor de suporte (SVM) e redes neurais convolucionais (CNN) foram aplicados para prever o desempenho do aluno. SVM é uma técnica de classificação binária com base em aprendizado de máquina supervisionado no campo da inteligência artificial (Drucker, Burges, Kaufman, Smola, & Vapnik,

1997). SVM básica leva um conjunto de dados de entrada para cada entrada dada e prevê que em duas classes possíveis formarão a saída, tornando-se um classificador linear binário não probabilística (Cortes & Vapnik, 1995). Há muitos benefícios para a aplicação SVM no processamento de dados educacionais, tais como a sua flexibilidade em relação às interações entre os parâmetros educacionais

de diferente fontes, tempos de formação de curta duração para criar novos modelos, e uma melhor sensibilidade para a modelagem de variáveis dependentes (Ifenthaler & Widanapathirana, 2014). Nos últimos anos, as redes neurais profundas, como a CNN, que emprega camadas de convolução e agrupamento para capturar dependências úteis entre atributos, obtiveram resultados notáveis em várias tarefas de classificação, tais como classificação de texto (Kim, 2014) e classificação de imagens (Krizhevsky, Sutskever, & Hinton, 2012). Neste estudo, foram treinados um simples CNN com uma camada de convolução com base em ambos os atributos estruturados e não estruturados apresentados na Tabela 2.

Para a identificação precoce de pelo - estudantes de risco, examinamos a capacidade preditiva do modelo na quinta (primeiro trimestre), sétimo e nono semana (antes de médio prazo) do semestre. sessões de classe antes da quinta semana foram excluídos porque não havia passado tempo suficiente para coletar dados suficientes. Classes após a nona semana foram excluídos porque o feedback resultante teria chegado tarde demais para ajudar os alunos a melhorar o seu desempenho antes do curso terminou.

Os parâmetros de avaliação, a retirada (como Equação 2), precisão (como Equação 3), e F - medir (como Equação 4), são utilizados para avaliar o desempenho. F - medida, como a média harmônica de precisão e recall, é amplamente utilizado em domínios tais como recuperação de informação, aprendizagem de máquina, análise de sentimentos, e muitos outros que envolvem a classificação binária.

$$\text{recordação} = \frac{\text{número de falha ou passado}}{\text{número total de falha ou passado}} \quad (2)$$

$$\text{Precisão} = \frac{\text{número de falha ou passado}}{\text{número total de falha ou passado}} \quad (3)$$

$$F - \text{a medida} = \frac{2 \times \text{Precisão} \times \text{recall}}{\text{Precisão} + \text{Lembre-se}} \quad (4)$$

4 | RESULTADOS E VISUALIZAÇÃO

4.1 | Resultados experimentais

4.1.1 | previsão emoção

O léxico supracitado - abordagem foi utilizada para determinar os tipos de emoção (positivo, negativo e neutro) do 2449 auto - comentários de avaliação. A verdade solo dos tipos de emoção para estes comentários são apresentados na Tabela 3. Comparando os tipos de emoção previstos pelo léxico - abordagem baseada ea verdade chão gerada por anotadores humanos mostra uma precisão de 0,76.

4.1.2 | previsão de desempenho acadêmico

Usamos técnicas de aprendizado de máquina e vários atributos significativos, incluindo dados estruturados e não estruturados, para avaliar cedo -

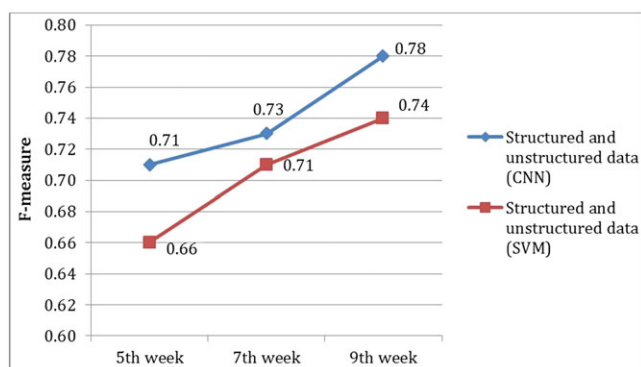


FIGURA 2 Cedo - encenar capacidade preditiva, resultado comparativo. CNN = redes neurais convolucionais; SVM = apoio máquina vetor [figura cor pode ser visto em wileyonlinelibrary.com]

TABELA 4 matriz de confusão das redes neurais convolucionais resultar na Semana 9

	prever	
	falha	Passar
Real		
Falhou	20	7
Passar	9	37

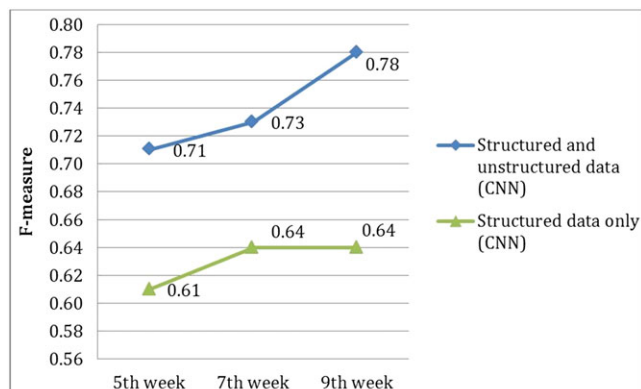


FIGURA 3 resultado comparativo da adição de dados não estruturados. CNN = redes neurais convolucionais; SVM = apoio máquina vetor [figura cor pode ser visto em wileyonlinelibrary.com]

encenar capacidade de previsão do desempenho do estudante nas semanas 5, 7 e 9 em um 19 - semana semestre. Como mostrado na Tabela 2, os dados estruturado inclui a frequência dos alunos, em - participação da classe, e conclusão lição de casa, ao passo que os dados não estruturados são tomadas a partir do conteúdo emocional dos comentários dos alunos avaliada manualmente por anotadores. Os dados para os 181 estudantes foram divididos aleatoriamente em formação e conjuntos de teste a uma proporção de 6:4 (108:73 estudantes).

Os resultados experimentais mostram o modelo proposto proporciona uma eficácia (representado pela F - medir) de 0,66 (SVM) e 0,71 (CNN) na Semana 5, melhorando a 0,71 / 0,73 na Semana 7, e 0,74 / 0,78

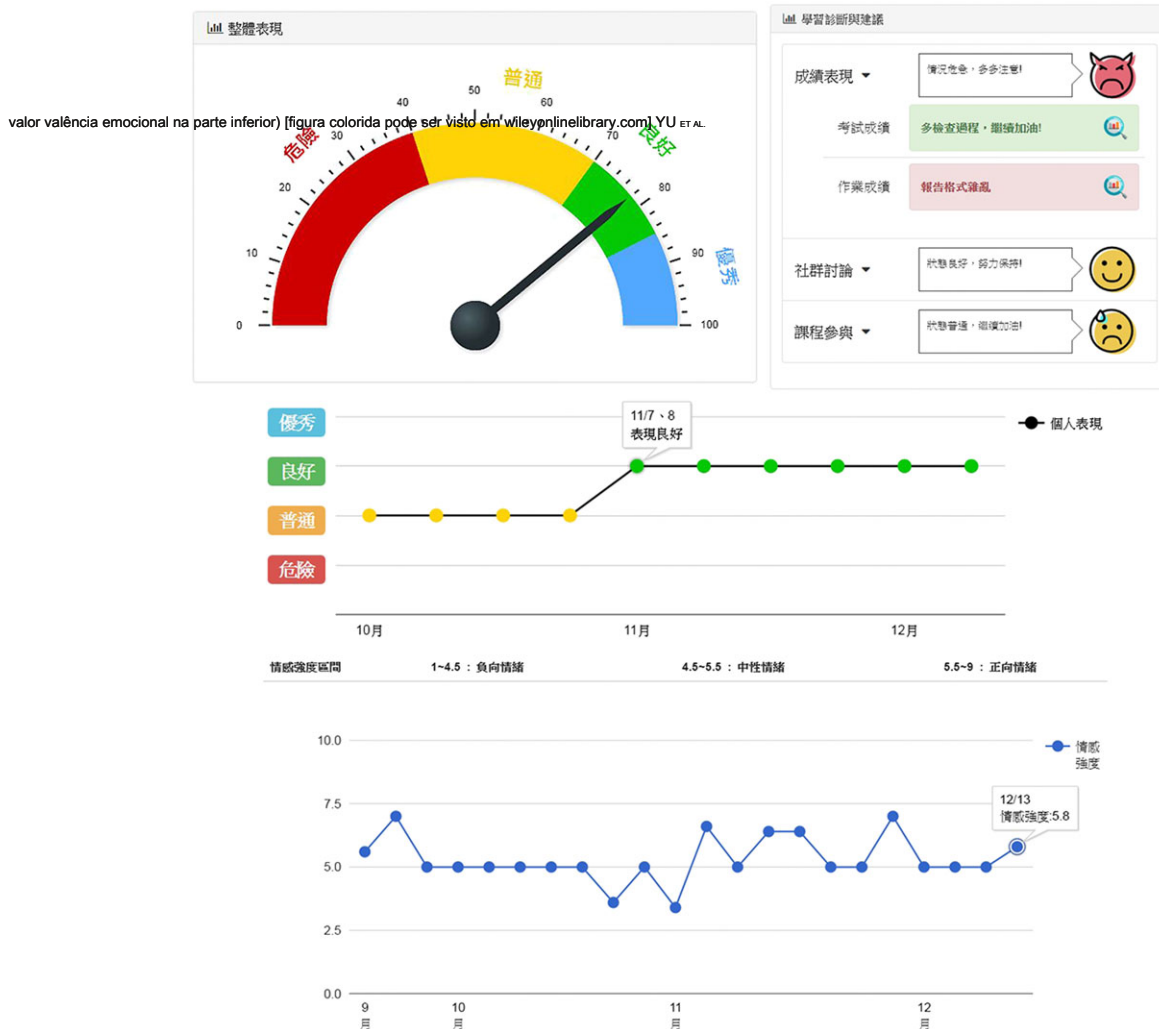
na semana 9. Os resultados na Figura 2 mostram que o modelo proposto apresentou o melhor eficácia na CNN, alcançar um F - medir o valor de 0,78 antes do médio prazo. A Tabela 4 mostra a matriz de confusão da CNN na Semana 9.

Para investigar a eficácia da análise de sentimento, nós examinamos as diferenças de F - medir com e sem dados não estruturados usando o algoritmo CNN. Os resultados experimentais mostram a F - medida sobe de 0,61 na Semana 5-0,64 nas semanas 7 e 9, utilizando apenas os dados estruturados. Resultados de comparação na Figura 3 mostra a melhoria de F - medir pela adição de dados não estruturados e

TABELA 5 Os resultados da experiência de CNN e SVM

		resultados da CNN F - medida (precisão)			resultados SVM F - medida (precisão)		
		5ª semana	7ª semana	9ª semana	5ª semana	7ª semana	9ª semana
somente os dados estruturados		0,61 (0,616)	0,64 (0,630)	0,64 (0,630)	0,49 (0,630)	0,66 (0,698)	0,60 (0,643)
dados estruturados e não estruturados		0,71 (0,712)	0,73 (0,739)	0,78 (0,780)	0,66 (0,657)	0,71 (0,712)	0,74 (0,739)
T teste	n	73	73	73	73	73	73
	t - valores	3.604	2.742	1,832	- 6,693	- 2.917	- 2.917
	p (bicaudal)	.000	.007	.069	.000	.004	.004

Nota. Os dados para os 181 estudantes foram divididos aleatoriamente em formação e conjuntos de teste a uma proporção de 6:4 (108:73 estudantes). CNN = redes neurais convolucionais; SVM = máquina de vetores de suporte.



A FIGURA 4 Dinâmica de diagnóstico e auto - regulada (DDS) do sistema (dashboard no canto superior esquerdo; relatório de diagnóstico e sugestão no canto superior direito; de status semanal no meio, e

demonstra a eficácia do uso de auto - comentários de avaliação na previsão precoce.

Um emparelhado, dois - atados t teste foi utilizado para comparar os resultados preditivos do uso de dados estruturados sozinho e com dados não estruturados. Tabela 5 lista os F - medida, precisão e nível de significância do t teste relacionado a CNN e SVM. Os resultados presentes são estatisticamente significativos ao nível de 0,05, excepto para a semana 9 da CNN ($p = .069$). Esta descoberta indica que o uso de emoções dos alunos extraído do auto desestruturado - comentários avaliadas melhorou significativamente a capacidade de previsão do classificador, especialmente para as fases anteriores (por exemplo, Semana 5 e 7). Uma razão possível é que, em tais fases anteriores, os dados estruturados são relativamente muito limitados para proporcionar suficiente

informação para a previsão; portanto, a comentários não estruturados têm um efeito mais pronunciado em melhorar o desempenho de previsão. Por outro lado, pela fase posterior (por exemplo, Semana 9), os dados estruturados suficientes tenham sido acumulada. reconhecimento atempado dos estados emocionais dos alunos pode ajudar educadores a compreender melhor o estado de aprendizagem de todos os membros da classe, o que lhes permite ajustar rapidamente suas estratégias pedagógicas para a eficácia máxima.

4.2 | visualização de informação

Um sistema de DDS é desenvolvida na base dos dados estruturados e não estruturados utilizados nas experiências. Como mostrado na Figura 4, o sistema de DDS proporciona um relatório de diagnóstico com base na visualizada fina - grained aprender informações de status. Os relatórios incluem quatro componentes: um painel fornecendo uma visão ampla avaliação; diagnóstico e desempenho - recomendações base para a actividade -

instrução específica; relatórios de status semanais; e os valores de valência emocional para estado emocional estudante após cada lição.

O painel usa cores diferentes para representar o estado aprendizagem dos alunos da seguinte forma: azul (excelente); verde (Boa); amarelo (justo); e vermelho (em - risco). Esta representação visual oferece aos alunos e instrutores com a aprendizagem pessoal informações de status em um relance. Diagnósticos e recomendações são fornecidos em três categorias: classes (incluindo exames e trabalhos de casa), a discussão sobre mídia social, e em - participação da classe. Cada categoria é representada por um ícone que os alunos podem clicar para receber recomendações e desempenho adicional - informações relacionadas a partir de assistentes de ensino ou professores. Por exemplo, na Figura 4, a recomendação de texto para " exames "

sugere que o estudante volte a verificar o processo de cálculo para questões matemáticas, e que por " dever de casa " sugere que o aluno melhorar o formato do relatório confuso. As recomendações para cada categoria de ajudar os alunos a compreender a sua avaliação de desempenho e para responder em conformidade. de status semanal é apresentado longitudinalmente em um gráfico de linha que pode ser usado para controlar as alterações no desempenho dos alunos ao longo do semestre. O gráfico valor valência emocional ajuda tendências de status trilha estudante de aprendizagem, com deslocamentos para baixo em estado emocional provocando um alerta. Tempo - auto relevante - comentários de avaliação pode ser acessado pelo direito - clicando em qualquer lugar ao longo da linha. Este gráfico pode ajudar os alunos a refletir sobre suas emoções relacionadas ao curso, a fim de melhorar o envolvimento, a conclusão lição de casa, e desempenho nos testes.

5 | Resultados e conclusões

Auto - comentários de avaliação, uma forma de dados não estruturados, fornecer informações valiosas que podem representar o estado de aprendizagem dos alunos ao longo da duração de um curso de instrução. emoções dos alunos, extraídas de tais comentários, refletem suas atitudes de aprendizagem em relação a lição, compreensão do conteúdo do curso, e dificuldades em se inclinando. Este estudo se aplica análise de sentimentos de auto - avaliação comenta para melhor identificar a - estudantes de risco. Os resultados experimentais demonstraram que a adição extraído auto informação sentimento fromstudents' - comentários de avaliação melhora significativamente a precisão da previsão inicial. Os resultados também indicam o início limitado - encenar capacidade preditiva de dados estruturados (por exemplo, a frequência, a realização trabalhos de casa, e em - participação da classe), devido à escassez de dados no início do curso. Assim, a aplicação de análise de sentimentos de dados não estruturados (por exemplo, auto - comentários de avaliação) pode melhorar significativamente a precisão do início - previsões de palco.

As principais contribuições deste estudo são duas. Em primeiro lugar, a identificação precoce de alunos que são propensos a falhar é fundamental para o sucesso de qualquer intervenção pedagógica destinada a mudar a trajetória de uma em - estudante risco. Jayaprakash et al. (2014) sugeriu que a probabilidade de melhoria diminui significativamente à medida que o tempo passa.

Aplicando dados não estruturados de emoções dos alunos para melhorar o início - precisão da previsão estágio pode permitir que o educador e educando para resolver a questão no tempo para resgatar nota final do aluno. Segundo, o sistema DDS proposta prevê um apoio valioso para ambos os educadores e estudantes. Pekrun et al. (2002) observou que as emoções positivas promover a auto - regulação entre os estudantes, enquanto que as emoções negativas levam a dependência de orientação externa. Os relatórios de diagnóstico e sugestões no sistema DDS fornecer informações úteis para a auto - regulada learningwith emoções positivas, eo gráfico de linhas do valor de valência fornece instrutores com informações importantes para a aplicação da orientação externa para alunos com emoções negativas. Nossos resultados permitem que os educadores para fornecer aos alunos reais - feedback em tempo e aumentar a autonomia do aluno.

Este estudo tem duas limitações. Primeiro, o estudo é baseado em uma pequena amostra de alunos de duas seções de um único curso em uma única universidade, e o trabalho adicional é necessário antes que os resultados podem ser generalizados para outros grupos. Em segundo lugar, o léxico - abordagem utilizada para a emoção previsão não leva em conta a negação (por exemplo, o estudante não é confuso). No entanto, apenas 5% (115?

2449) de todos os comentários no conjunto experimental de dados negação contido. O trabalho futuro incidirá sobre a aplicação do modelo de conjuntos de dados maiores a partir de várias universidades e tipos de curso e melhorar a eficácia ea usabilidade do sistema DDS para estudantes e educadores.

AGRADECIMENTOS

Este trabalho foi financiado pelo Ministério da Ciência e Tecnologia, Taiwan, ROC, sob Grants MAIORIA 105 - 2221 - E - 155 - 059 - MY2 e MAIS 105 - 2218 - E - 006 - 028. Os autores gostariam de agradecer aos revisores anônimos e as cadeiras da área por seus comentários construtivos.

ORCID

LC Yu <http://orcid.org/0000-0003-1443-4347>

CW Lee <http://orcid.org/0000-0001-9098-0660>

REFERÊNCIAS

- Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2015). aprendizagem prevendo - emoções relacionadas de textual via Twitter gabarito sala de aula dos alunos. Banea, C.,
- Mihalcea, R., & Wiebe, J. (2013). Portando subjetividade multilingue recursos entre idiomas. *IEEE Transactions on Affective Computing*, 4 (2), 211 - 225.
- Bayer, J., Bydzovská, H., Géryk, J., Obsivac, T., & Popelinsky, L. (2012). prevendo queda - para fora do comportamento social dos alunos. *Internacional para a Educação de Dados Sociedade Mineira*.
- Blikstein, P., & Worsley, M. (2016). analytics aprendizagem multimodais e mineração de dados educação: Usando tecnologias computacionais para medir tarefas de aprendizagem complexos. *Journal of Learning Analytics*, 3 (2), 220 - 238.
- Calvo, RA, e D'Mello, S. (2010). Afectam a detecção: Uma interdisciplinar revisão de modelos, métodos e suas aplicações. *IEEE Transactions on Affective Computing*, 1 (1), 18 - 37.
- Cassady, JC, & Johnson, RE (2002). a ansiedade do teste cognitivo e acadêmico desempenho. *Psicologia da Educação Contemporânea*, 27 (2), 270 - 295.
- Cohen, J. (1960). Um coeficiente de concordância para escalas nominais. *Educacional e psicológico de medição*, 20 (1), 37 - 46.
- Cortes, C., & Vapnik, V. (1995). Apoio, suporte - redes vetor. *Machine Learning*, 20 (3), 273 - 297.
- D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., Branco, H., ... Picard, R. (2008). AutoTutor detecta e responde aos alunos afetivos e estados cognitivos. Trabalho apresentado no Workshop sobre questões emocionais e cognitivos na Conferência Internacional sobre Sistemas Tutores Inteligentes.
- Drucker, H., Burges, CJ, Kaufman, L., Smola, A., & Vapnik, V. (1997). máquinas de regressão vector de apoio. *Avanços em Sistemas de Processamento de Informações Neurais*, 9, 155 - 161.
- Ekman, P. (1992). Um argumento para emoções básicas. *Cognição e emoção*, 6 (3 - 4), 169 - 200.
- Feldman, R. (2013). Técnicas e aplicações para análise de sentimentos. *Communications of the ACM*, 56 (4), 82 - 89.
- Hembree, R. (1988). Correlates, faz com que, de efeitos, e o tratamento de teste ansiedade. *Revisão de Pesquisas Educacionais*, 58 (1), 47 - 77.
- Ilfenthaler, D., & Widanapathirana, C. (2014). Desenvolvimento e validação de um quadro de análise de aprendizagem: Dois estudos de caso utilizando máquinas de vetor de suporte. *Tecnologia, Conhecimento e Aprendizagem*, 19 (1 - 2), 221 - 240.
- Jayaprakash, SM, Moody, EW, Lauria, EJ, Regan, JR, & Baron, JD (2014). alerta precoce de academicamente na - estudantes de risco: An open iniciativa analytics origem. *Journal of Learning Analytics*, 1 (1), 6 - 47.
- Kim, Y. (2014). redes neurais convolucionais para a classificação sentença. *arXiv preprint arXiv: 1408.5882*.
- Kiritchenko, S., & Mohammad, SM (2016). Capturando bem confiável - granulação associações sentimento de crowdsourcing e melhor - pior escala. Trabalho apresentado na revista *Proceedings, NAACL - HLT*.
- Kiritchenko, S., Mohammad, SM, & Salameh, M. (2016). SemEval - 2016 Tarefa 7: Determinar a intensidade sentimento de Inglês e frases em árabe. *Proceedings of SemEval*, 42 - 51.
- Krizhevsky, A., Sutskever, I., & Hinton, GE (2012). classificação IMAGENet com redes neurais convolucionais profundas. Trabalho apresentado no *Avanços em sistemas de processamento de informações neurais*. Lang, PJ (1980). Tratamento comportamental e bio - avaliação comportamental: aplicações informáticas.
- Li, J., Ott, M., Cardie, C., & Hovy, EH (2014). Para uma regra geral para a identificação enganosa opinião spam. Trabalho apresentado no *ACL*, 1.
- Luo, J., Sorour, SE, Goda, K., & mina, T. (2015). Prevendo grau estudante baseado na livre - estilo comenta using Word2vec e ANN considerando resultados de previsão obtidos em aulas consecutivas. *Internacional Educacional Dados Sociedade Mineira*.
- Mishne, G., & De Rijke, M. (2006). Capturando os níveis de humor globais usando blogue Postagens. Trabalho apresentado no simpósio primavera AAAI: Computacional abordagens para analisar weblogs.
- Nguyen, T., Phung, D., Dao, B., Venkatesh, S., e Berk, M. (2014). Afetivo e análise de conteúdo das comunidades depressão on-line. *IEEE Transações em Affective Computing*, 5 (3), 217 - 226.
- Pekrun, R., Goetz, T., Titz, W., & Perry, RP (2002). emoções acadêmicos em auto dos alunos - aprendizagem regulada e realização: Um programa de pesquisa qualitativa e quantitativa. *Psicólogo educacional*, 37 (2), 91 - 105.
- Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutopoulos, I., & Manandhar, S. (2014). Semeval - 2014 tarefa 4: análise de sentimento baseado Aspect. *Proceedings of SemEval*, 27 - 35.
- Pritchard, ME, & Wilson, GS (2003). Usando fatores emocionais e sociais para prever o sucesso do aluno. *Jornal do Colégio Desenvolvimento Student*, 44 (1), 18 - 28.
- Romero, C., & Ventura, S. (2010). mineração de dados educacional: Uma revisão da Estado da arte. *IEEE Transactions on Systems, Homem, e Cibernética, Parte C (Aplicações e dicas)*, 40 (6), 601 - 618.
- Rosenthal, S., Nakov, P., Kiritchenko, S., Mohammad, SM, Ritter, A., & Stoyanov, V. (2015). Semeval - 2015 tarefa 10: análise de sentimentos no Twitter. Trabalho apresentado na revista *Proceedings, da Oficina Internacional 9 em Semantic Avaliação (SemEval 2015)*. Saif, H., Fernandez, M., Ele, Y., e Alani, H. (2013). conjuntos de dados de avaliação para Twitter análise de sentimentos: um levantamento e um novo conjunto de dados, o STS - Ouro.
- Santana, MA, Costa, EB, Fonseca, B., Rego, J., & de Araújo, FF (2017). Avaliar a eficácia das técnicas de mineração de dados educacionais para a previsão antecipada de insucesso escolar dos alunos em cursos introdutórios de programação. *Computadores no comportamento humano*. <https://doi.org/10.1016/j.chb.2017.01.047>
- Schouten, K., & Frasinicar, F. (2016). Pesquisa sobre os aspectos - sentimento nível análise. *IEEE Transactions on conhecimentos e dados de engenharia*, 28 (3), 813 - 830.
- Shahiri, AM, Husain, W., e Rashid, N. a. A. (2015). Uma revisão na previsão técnicas de mineração de dados de desempenho usando aluno. *Procedia Ciência da Computação*, 72, 414 - 422.
- Simon, HA (1996). As ciências do artificial (3ª ed, orig. publ. de 1969 ed.). Cambridge, MA: MIT Press.
- Smith, VC, Lange, A., & Huston, DR (2012). modelagem preditiva para prever os resultados dos alunos e dirigir intervenções eficazes em cursos universitários comunidade on-line. *Journal of assíncronos de aprendizagem Networks*, 16 (3), 51 - 61.
- Sorour, SE, Goda, K., & mina, T. (2015). Avaliação da eficácia da Tempo - comentários da série, utilizando técnicas de aprendizado de máquina. *Journal of Information Processing*, 23 (6), 784 - 794.
- Sorour, SE, mina, T. Goda, K., & Hirokawa, S. (2015). Um modelo preditivo para avaliar o desempenho dos alunos. *Journal of Information Processing*, 23 (2), 192 - 201.
- Spence, R. (2001). visualização de informação (vol. 1). Springer. Xu, R., Gui, L., Xu, J., Lu, Q., e Wong, K. - F. (2015). Cruz lingual extração titular opinião com base em múltiplos - kernel do SVMs e aprendizagem de transferência. *World Wide Web*, 18 (2), 299 - 316.
- Yu, L. - C., Lee, L. - H., Hao, S., Wang, J., He, Y., Hu, J., ... Zhang, X. (2016). Construir recursos afetivos chineses em Valence - dimensões excitação. Trabalho apresentado na revista *Proceedings, NAACL - HLT*.

Como citar este artigo: Yu LC, Lee CW, Pan-HI, et al. Melhorar a previsão inicial de insucesso escolar utilizando análise de sentimentos na auto - comentários avaliadas. *J Comput Assist Aprender*. 2018; 34: 358 - 365. <https://doi.org/10.1111/jcal.12247>