

Método Supervisionado para Identificação de Dúvidas em Fóruns Educacionais

Vitor B. Rolim¹, Rafael Ferreira¹, Evandro Costa²

¹DEINFO– Universidade Federal Rural de Pernambuco (UFRPE)
Caixa Postal s/n – 52171-900 – Recife – PE – Brasil

²Instituto de Computação – Universidade Federal Rural de Alagoas (UFAL)
Caixa Postal s/n – 57072-900 – Maceió – AL – Brasil

{victor.b.rolim,rafaelflmello,ebcosta}@gmail.com

Abstract. *Educational forum provides interaction between students and teachers, however the teacher's work may be hampered by the large number of posts. This paper proposes a method to identify questions in student's posts. This method involves two steps: classify the post and extract the subject of the post. Knowing the subject of the student's question posts, an auxiliary study material could be recommended.*

Resumo. *O fóruns educacionais proporciona a interação entre alunos e professores, porém o trabalho do professor pode ser dificultado pelo grande número de postagens. Este trabalho propõe um método para identificar dúvidas nas postagens dos alunos. Esse método envolve duas etapas: classificar a postagem e extrair o assunto da postagem. Conhecendo o assunto da postagem de dúvida do aluno, um material de estudo auxiliar pode ser recomendado.*

1. Introdução

Com o crescente uso da tecnologia como ferramenta de apoio educacional, o uso de Ambientes Virtuais de Aprendizagem (AVA) [Dillenbourg e Schneider 2002] tem aumentado nos últimos anos. Estes ambientes disponibilizam várias ferramentas para melhorar a interação entre professores e alunos, onde alguns exemplos são: fórum, blog, wiki, redes sociais, entre outros.

Estas ferramentas possuem um grande potencial para gerar conteúdo, o que pode ser usado para auxiliar no processo ensino-aprendizagem. Porém, devido à grande quantidade de interações entre os alunos e o professor, torna-se difícil para o professor proporcionar um auxílio personalizado para um determinado aluno, por isso é importante que os AVAs ofereçam meios de acompanhamento direto e indireto para garantir o aprendizado do aluno [Akyuz and Kurt 2010].

O acompanhamento direto é aquele realizado sob a supervisão do professor ou tutor. Para isso, o curso tem um plano de ensino e cronograma de atividades que são acompanhados de perto pelos professores e tutores. Então todo material e discussões disponibilizados no AVA é verificado manualmente para que as dúvidas dos alunos

sejam resolvidas e seus progressos computados. Devido à grande quantidade de informação, esse tipo de acompanhamento se torna por vezes inviável de ser realizado.

Para amenizar essa situação, é necessário também realizar o acompanhamento indireto, que é o acompanhamento sem a participação direta do professor. Para isso, é importante ter sistemas automatizados que possam auxiliar o professor nessa atividade.

A ferramenta assíncrona de fórum tem uma característica importante, é nela que os alunos postam dúvidas, comentários sobre a disciplina, outras fontes de assunto, possíveis respostas para questões levantadas pelo professor, entre outros. Cheng *et al.* [Cheng *et al.* 2011] realizou um estudo que mostra a efetividade da ferramenta de fórum para melhorar a performance de estudantes em um AVA.

Em um acompanhamento direto as postagens de dúvidas devem receber uma solução. Por outro lado, uma resposta pode ser usada para perceber o progresso do aluno, com isso o professor pode pontuá-lo ou pode utilizar o aluno como propagador do assunto entre os colegas [Kim 2013].

Para realizar o acompanhamento indireto de um fórum educacional de forma eficiente é importante a utilização de sistemas automáticos [Rolim *et al.* 2014, Gerosa *et al.* 2003]. Por exemplo, um sistema para recomendar automaticamente materiais direcionados para as dúvidas de cada aluno [Mohamad e Tasir 2013], ou um sistema que indicasse ao professor as postagens mais relevantes de acordo com o assunto abordado.

Este trabalho terá como principal objetivo propor um sistema para acompanhamento indireto de fóruns educacionais. Para isso, o primeiro passo é realizar a identificação do tipo de postagem feita no fórum educacional. O segundo passo é extrair o assunto contido na postagem. Uma vez identificado o tipo de postagem e o assunto que ela possui, se a postagem for uma dúvida, direciona conteúdos relacionados ao assunto extraído. Este sistema identifica automaticamente postagens que correspondem a uma dúvida, extrai o assunto abordado por uma determinada postagem e recomenda um material de estudo que possa ajudar ao aluno.

2. Trabalhos relacionados

Gerosa *et al.* (2003) apresenta o AulaNet, um ambiente de interação entre alunos e professores, e demonstra o encadeamento e a organização das postagens através de hierarquia. Essa hierarquia das postagens é derivada da categorização de cada uma delas, as categorias podem ser: seminário, questão, argumentação, contra argumentação, esclarecimento. Esta categorização, implementada pelo próprio AulaNet, ajuda a observar os relacionamentos entre as postagens, dando assim subsídios aos professores para que possam coordenar eficazmente essas discussões, e avaliar o desenvolvimento da turma.

Ravi e Kim (2007) propõe uma forma de identificar as interações dos alunos nos fóruns educacionais. Ele classifica individualmente as postagens dos alunos como “atos de discurso”, que podem ser: complemento, informação, correção, elaboração, pergunta, resposta. Para a categorização são utilizadas técnicas de análise, *N-grams* e SVM para encontrar postagens que sejam categorizadas como pergunta ou resposta, e assim poder auxiliar o professor em encontrar perguntas sem respostas.

Lin *et al.* (2009) propõe um sistema de classificação de gênero das postagens dos fóruns de discussão em ambientes educacionais. Utilizando a frequência das palavras como características para o sistema, e aplicando árvore de decisão para classificação. Os gêneros que este sistema tenta identificar são: anúncios (esclarece dúvidas dos outros), perguntas, interpretação, conflito (opinião conflitante), afirmação, e outros (mensagens variadas que são difíceis de categorizar).

Oliveira Júnior *et al.* (2011) apresenta uma ferramenta de classificação automática de postagens em fóruns educacionais. As postagens podem ser classificadas como positivas ou negativas, utilizando algoritmos de aprendizado de máquina, como *Naive Bayes* e SVM. Essa classificação das postagens como positivas (expressam respostas, comentários pertinentes entre outros) e negativas (expressam dúvidas, conteúdo indevido entre outros) auxiliam o professor no momento de fornecer uma atenção maior a um determinado assunto.

3. Metodologia

3.1. Classificação das postagens

Esta etapa tem por objetivo identificar dúvidas nas postagens de fóruns educacionais. Para isso foi utilizado técnicas de aprendizado de máquina para classificá-las. Essas postagens podem ser uma dúvida, uma resposta ou um comentário neutro.

O desenvolvimento desta etapa se dividiu em cinco fases: coleta da base de dados, pré-processamento, extração das características, classificação e análise de resultados.

3.1.1. Coleta da base de dados

Durante o processo de desenvolvimento desta fase de classificação, foram utilizadas 2 bases de dados diferentes.

A base de dados 1 (BD1) foi retirada do AVA da Universidade Federal de Alagoas (UFAL), e as postagens foram coletadas de várias disciplinas do curso à distância de bacharelado em Sistemas de Informação. Já a base de dados 2 (BD2), foi retirada também do AVA da UFAL, porém, das disciplinas de Algoritmos e Estruturas de Dados I e II, do mesmo curso.

A base de dados 1 é composta por 490 postagens, sendo: 198 de postagens contendo dúvida; 104 postagens que são comentários neutros; 188 de postagens contendo respostas. Já a base de dados 2 é composta por 600 postagens, com a seguinte distribuição: 200 postagens contendo dúvida; 200 postagens que são comentários neutros; 200 postagens contendo respostas.

3.1.2. Pré-processamento

A partir da postagem do fórum é realizada a etapa de pré-processamento do texto. Para esta etapa, duas técnicas foram empregadas: a remoção de *stopwords* e o *stemming* [Hotho *et al.* 2005].

As *stopwords* são as palavras que com pouca representatividade para os textos, por exemplo: artigos e pronomes. A palavra *stemming* deriva da palavra inglesa *stem*, que significa “radical”, por sua vez, técnica homônima transforma a palavra na sua forma radical, removendo sufixos.

Além dessas técnicas, também foi realizada a remoção de termos que tenham uma baixa frequência absoluta em relação à base de dados. Foram estabelecidos alguns limites de corte, para que palavras com o valor da frequência abaixo do limite sejam descartadas. Essa remoção é feita porque os termos com baixa frequência normalmente têm pouca influência na identificação das classes definidas.

3.1.3. Extração das Características

Após o pré-processamento, foi realizada a extração das características que serão os atributos das técnicas de classificação. Foram definidos dois conjuntos de características. Entre as características utilizadas no primeiro conjunto estão: (i) frequência das palavras da classe Dúvida; (ii) frequência das palavras da classe Neutra; (iii) frequência das palavras da classe Resposta; (iv) número de interrogações.

Foram realizados testes usando conjuntos de características diferentes, em busca do melhor resultado. As características utilizadas do segundo conjunto são: (i) somatório do TF-IDF das palavras pertencentes a classe Dúvida; (ii) somatório do TF-IDF das palavras pertencentes a classe Neutra; (iii) somatório do TF-IDF das palavras pertencentes a classe Resposta; (iv) número de interrogações.

O TF-IDF (*Term Frequency – Inverse Document Frequency*) [Salton e Buckley 1988] de cada palavra é calculado e multiplicado pelo número de aparições da palavra, antes do somatório. Uma vez obtidos esses atributos, eles são disponibilizados como entrada para as técnicas de classificação.

O uso da característica de número de interrogações por postagem também é um parâmetro interessante, pois o caractere de interrogação está bastante frequente em postagens da classe de dúvida, embora seja possível encontrar a interrogação em postagens de outras classes.

3.1.4. Classificação e Métricas de Avaliação

Neste trabalho utilizamos a ferramenta Weka (*Waikato Environment for Knowledge Analysis*) [Frank *et al.* 2010], versão 3.6.10 para realizar a classificação e avaliação dos resultados.

Para avaliar os classificadores foram usadas as seguintes métricas: Precisão, Cobertura e *F-Measure* [Friedman *et al.* 1997]. Precisão avalia a quantidade de instâncias que foram classificadas corretamente. Cobertura avalia a porcentagem instâncias de uma determinada classe que não foi classificada como pertencente a essa classe. *F-Measure* pode ser interpretada como uma média harmônica da precisão e da cobertura.

3.2. Extração do assunto das postagens

Esta etapa é responsável por identificar qual é o assunto contido na postagem do aluno. Para compreender melhor a metodologia deste trabalho podemos observar a Figura 1.

O fluxo deste projeto foi dividido em duas etapas, a primeira etapa recebe como entrada o conteúdo programático de uma disciplina provido pelo professor da mesma. Logo após receber essa entrada, todos os assuntos descritos no conteúdo programático são buscados no Wikipédia, e um banco de palavras é criado para cada um dos assuntos, cada palavra possui um peso associado.

Em seguida a postagem é submetida as técnicas de pré-processamento. A postagem então é comparada com os bancos de palavras, e verifica-se qual tem maior valor pelo somatório dos pesos das palavras. O assunto em que o somatório dos pesos for maior, é definido como assunto da postagem.

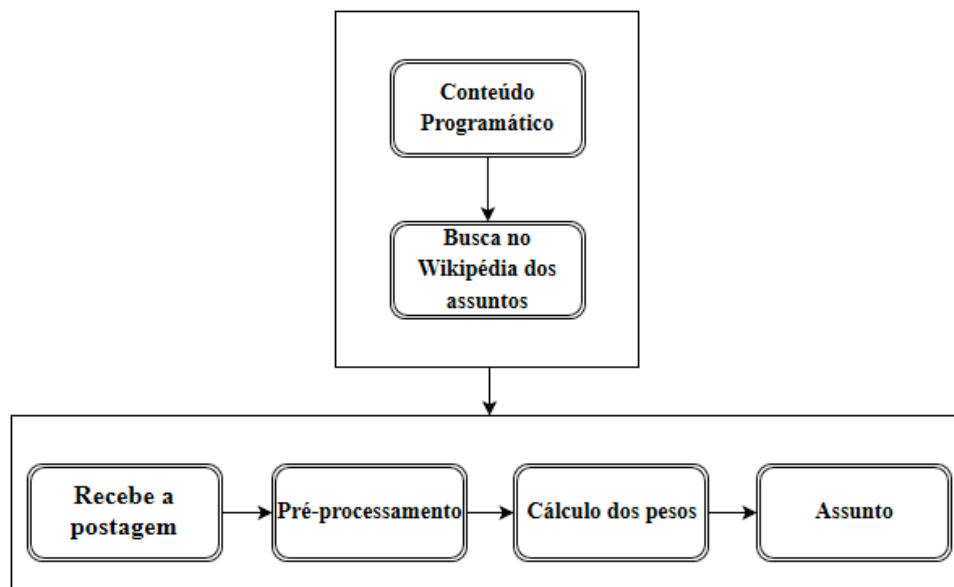


Figura 1 - Metodologia do sistema de extração de assunto.

3.2.1. Busca dos assuntos no Wikipédia

Quando uma disciplina é cadastrada no AVA, é necessário também que seja adicionado o plano de trabalho da disciplina, que aborda várias informações sobre ela. Dentre essas informações, existe o conteúdo programático, nele estão todos os assuntos que serão vistos no decorrer da disciplina.

Foi criado uma função que reconhece e extrai o conteúdo programático do plano de trabalho da disciplina. Em seguida outra função ainda organiza hierarquicamente os assuntos, a Figura 2 mostra um exemplo de hierarquia dos assuntos.

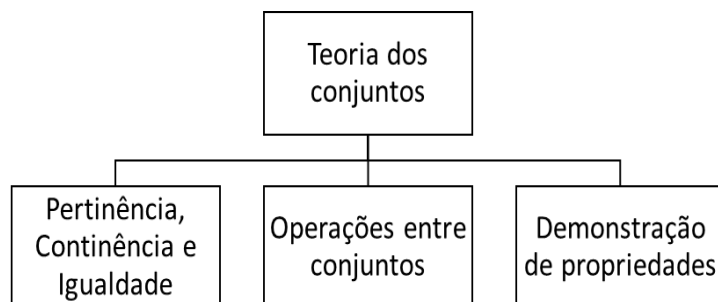


Figura 2 - Exemplo da hierarquia dos assuntos.

Para cada assunto é feito uma consulta no Wikipédia de cada um deles. Todos os *hiperlinks* e as palavras em negrito são adicionados no banco de termos de seu respectivo assunto, juntamente com seu peso. O peso é calculado pela equação do TF (*Term frequency*), que é razão entre a quantidade de aparições do termo e a quantidade total de termos no texto do da página consultada no Wikipédia.

3.2.2. Base de dados

Para a etapa de extração de assunto foram utilizadas duas bases de dados, a base de dados 3 (BD3) possui 387 postagens, foi desenvolvida na disciplina de mineração de texto, do curso de bacharelado em ciência da computação da UFRPE. Os alunos dessa disciplina tiveram que escrever postagens que apresentassem dúvidas ou respostas com relação aos assuntos da disciplina de matemática discreta.

A BD3 foi distribuída segundo o assunto que representavam na disciplina proposta, e tem a seguinte distribuição: 122 postagens de “Relações, funções e sequências”; 118 postagens de “Lógica”; 119 postagens de “Teoria dos conjuntos”; 38 postagens de “Teoria dos números”. A outra base de dados é a BD2 que também foi utilizada na etapa de classificação, mais detalhes sobre essa base de dados pode ser encontrado na Seção 3.1.1.

Estas bases serão utilizadas para que se possa obter a acurácia do algoritmo de extração de assunto. Com esse intuito a base de dados 3 foi criada, por isso ela obedece uma distribuição diferente, essa distribuição facilita a análise dos resultados. A base de dados 2 também será usada para avaliar a taxa de acerto do algoritmo, porém, como ela possui uma distribuição diferente, esta avaliação é feita manualmente, verificando cada postagem e o assunto indicado.

3.2.3. Pré-processamento

Para este trabalho foram aplicadas duas técnicas de pré-processamento textual, a remoção de *stopwords* e o *stemming*.

3.2.4. Cálculo dos pesos

Como já mencionado, o cálculo dos pesos das postagens nos fóruns vai ser baseado nos bancos de termos extraídos da Wikipédia. O processo funciona da seguinte forma:

1. A postagem vai ser analisada para cada assunto do conteúdo programático, que contém um banco de termos. Para cada assunto:
 - a. Cada termo contido na postagem é comparado com os bancos de termos. Caso o termo esteja contido no banco o peso do termo é somado.
 - b. A soma final de cada banco é armazenada.
2. Ao final do processamento de todos os assuntos o somatório que possuir maior valor é definido como assunto da postagem.

3.3. Recomendação do material de estudo auxiliar

Esta é a última etapa do sistema, nesta etapa, é dado como entrada o assunto extraído da postagem assim como visto na etapa anterior, e é buscado um vídeo do Youtube que melhor atenda o assunto. O vídeo encontrado é sugerido ao aluno.

Após a etapa de extração do assunto da postagem, o assunto é passado como argumento para a função responsável por buscar vídeos no Youtube. Essa função retorna o link do vídeo melhor ranqueado, o algoritmo do Youtube por padrão ordena a lista de vídeos por relevância. Ao obter este link, ele é recomendado ao aluno, servindo assim como auxílio para que o aluno possa sanar sua dúvida.

4. Experimentos

4.1. Classificação das postagens

Para avaliarmos qual seria a melhor técnica de classificação para atingir o objetivo apresentado. As técnicas utilizadas foram: Classificador Bayesiano (NaiveBayes), Árvore de decisão (J48), Rede neural (MultilayerPerceptron).

Para a realização deste primeiro teste, não foi aplicado nenhum tipo de pré-processamento nas postagens, as características selecionadas foram as frequências de palavras de cada classe (Dúvida, Neutra, Resposta), e o limite de corte foi deixado em 0. Podemos observar o resultado na Tabela 1. A avaliação da classificação será feita usando a métrica *F-measure*.

Tabela 1 - Média das medidas-F das técnicas de classificação

<i>Técnicas de classificação</i>	<i>Média da F-measure</i>	
	<i>BD1</i>	<i>BD2</i>
NaiveBayes	0.526	0.701
J48	0.885	0.882
MultilayerPerceptron	0.962	0.972

Ao analisarmos os resultados, chega-se a conclusão que a rede neural MLP é a mais indicada para a classificação de postagens de fóruns educacionais, dentre as três técnicas analisadas, alcançando os valores de 0.962 e 0.972 para BD1 e BD2 respectivamente. A conclusão de ter um bom classificador para o problema impacta na possibilidade de identificar automaticamente o tipo de postagem do fórum, possibilitando ser integrado a um sistema que ajude o aluno quando houver postagem de dúvida. O resultado confirma a conclusão de Rolim *et al.* [Rolim *et al.* 2014], que aponta o MLP como melhor algoritmo classificador para a classificação de postagens.

Com o melhor algoritmo definido, resta definir quais serão as características que devem ser selecionadas, quais técnicas de pré-processamento devem ser empregadas, e qual limite de corte deve ser estabelecido, a fim de que possa ser encontrado o melhor cenário possível.

Como o principal objetivo do sistema é encontrar dúvidas, o melhor resultado, será o que possuir o maior valor da *F-measure* da classe Dúvida.

4.1.1. Usando frequência das palavras

Ao utilizar como características as frequências das palavras foi definido 3 cenários diferentes, são eles: (i) Cenário 1: Classificação sem técnicas de pré-processamento; (ii) Cenário 2: Classificação usando remoção de *stopwords*; (iii) Cenário 3: Classificação usando remoção de *stopwords* e *stemming*.

Os limites de cortes utilizados foram definidos com base nos limites utilizados em Rolim *et al.* [Rolim *et al.* 2014] e variaram entre 0, 5 e 10. Podemos observar os valores da *F-measure* alcançados para cada classe de acordo com a base de dados utilizadas na Tabela 2.

O cenário 1 e o cenário 2 apresentaram os maiores valores da *F-measure* da classe dúvida tanto na BD1 quanto na BD2 quando o limite de corte estava em 0.

Tabela 2 - F-measure das classes de acordo com a base de dados (Frequência).

		Limite de corte 0			Limite de corte 5			Limite de corte 10		
		Dúvida	Neutra	Reposta	Dúvida	Neutra	Reposta	Dúvida	Neutra	Reposta
BD1	Cenário 1	0.96	0.935	0.979	0.738	0.708	0.774	0.728	0.623	0.745
	Cenário 2	0.963	0.953	0.966	0.793	0.768	0.801	0.736	0.634	0.71
	Cenário 3	0.892	0.825	0.922	0.779	0.777	0.806	0.77	0.713	0.782
BD2	Cenário 1	0.977	0.961	0.977	0.915	0.844	0.835	0.821	0.797	0.827
	Cenário 2	0.962	0.933	0.955	0.906	0.859	0.872	0.891	0.823	0.834
	Cenário 3	0.872	0.874	0.895	0.879	0.809	0.854	0.842	0.785	0.796

4.1.2. Usando TF-ITF

Assim como na Seção 4.1.1 foram definidos 3 cenários diferentes: (i) Cenário 4: Classificação sem técnicas de pré-processamento; (ii) Cenário 5: Classificação usando remoção de stopwords; (iii) Cenário 6: Classificação usando remoção de stopwords e *stemming*.

O limite de corte, variou entre 0, 0.00109 e 0.00247. Os limites de cortes utilizados foram definidos após o teste de vários outros valores, os valores adotados foram o que apresentaram os melhores resultados. A Tabela 3 mostra os valores da *F-measure* para cada classe nas duas bases de dados.

Tabela 3 - F-measure das classes de acordo com a base de dados (TF-IDF).

		Limite de corte 0			Limite de corte 0.00109			Limite de corte 0.00247		
		Dúvida	Neutra	Reposta	Dúvida	Neutra	Reposta	Dúvida	Neutra	Reposta
BD1	Cenário 4	0.832	0.48	0.636	0.854	0.819	0.795	0.834	0.705	0.748
	Cenário 5	0.832	0.879	0.78	0.911	0.898	0.87	0.855	0.763	0.788
	Cenário 6	0.823	0.667	0.711	0.908	0.916	0.87	0.868	0.849	0.823
BD2	Cenário 4	0.938	0.908	0.895	0.885	0.817	0.777	0.846	0.655	0.585
	Cenário 5	0.85	0.549	0.536	0.941	0.899	0.895	0.893	0.822	0.8
	Cenário 6	0.843	0.639	0.611	0.926	0.885	0.88	0.882	0.793	0.801

Ao observarmos os resultados apresentados, notamos que os dois primeiros cenários (usando frequência de palavras) obtiveram os melhores resultados, tanto na BD1 quanto na BD2, com os valores de 0.96 e 0.963 para a BD1, e 0.977 e 0.962 para a BD2. Podemos observar nas Tabelas 2 e 3 que os cenários 1 e 2 são os que apresentam os maiores valores da *F-measure* em ambas as bases de dados.

O cenário 1 faz uso da frequência das palavras das classes como característica, e não faz uso de técnicas de pré-processamento. Os valores apresentados na figura 1 para o cenário 1, foram alcançados quando o limite de corte estava em 0.

O cenário 2, faz uso da frequência das palavras das classes, e faz a remoção de *stopwords* como técnica de pré-processamento. Assim como o cenário 1, os melhores valores foram atingidos quando o limite de corte estava em 0.

Se calcularmos a média dos valores atingidos na BD1 e BD2 do cenário 1 e 2, chegamos aos valores de 0.968 e 0.962 respectivamente. O valor da média do cenário 1 é levemente superior à média do cenário 2. Portanto, o cenário 1 é o mais adequado para ser utilizado na etapa de classificação das postagens.

4.2. Extração do assunto das postagens

Como explicado na Seção 3.2, foram utilizadas duas bases de dados (BD2 e BD3) com o intuito de avaliar a taxa de acerto do algoritmo de extração de assunto. Para os testes realizados, foi necessário também o conteúdo programático dos cursos das quais as

bases estão relacionadas, no caso, algoritmos e estrutura de dados para a BD2, e matemática discreta para a BD3.

Foi possível extrair corretamente o assunto de 66,66% das postagens. Contudo, identificamos que algumas postagens dificultam a extração do seu assunto corretamente, tais como: “Professor, qual seria a melhor forma para resolver a primeira questão da lista?”; “Que horas abre o laboratório?”; “Qual será a data da prova?”.

Esses tipos de postagens dificultam a extração do seu assunto, sabendo que para uma extração correta, precisamos que o assunto esteja relacionado com algum assunto no conteúdo programático da disciplina. Realizamos outro teste removendo esses tipos de postagens, a taxa de acerto subiu para 90,58%. Este resultado mostra que a eliminação das postagens que não são relacionadas a algum assunto melhora a extração.

O resultado do texto no BD3 pode ser observado na Tabela 4. A média da taxa de acerto obtida foi de 76,1%. Analisando as partes que compõem este resultado, observamos que os resultados individuais foram diferentes, um exemplo dessa diferença pode ser visto pela taxa de acerto de “Relações, funções e sequências”, 62,3%, e pela taxa de “Lógica” de 86,44%.

Tabela 4 - Taxas de acerto dos assuntos da base de dados 1.

<i>Assunto</i>	<i>Taxa de acerto</i>
Teoria dos conjuntos	71,43%
Teoria dos números	84,21%
Relações, funções e sequências	62,3%
Lógica	86,44%
Média	76,1%

Esses resultados nos mostram que mesmo com assuntos do mesmo universo, no caso, matemática discreta, houve uma diferença considerável entre os resultados. Essa diferença pode ser creditada a como o conteúdo programático da disciplina é descrito.

6. Conclusão

Com a grande disseminação da educação a distância os ambientes virtuais de aprendizagem estão sendo largamente usados. Os fóruns têm uma característica importante, nele os alunos postam dúvidas e possíveis respostas para questões levantadas pelo professor.

Com isso, esta ferramenta produz um conteúdo bastante valioso para o curso. Contudo, devido à grande quantidade de alunos normalmente inscritos em cursos com plataformas online, torna-se difícil realizar o acompanhamento direcionado para cada aluno de forma personalizada.

Diante dessas considerações a proposta deste trabalho, foi a criação de uma nova abordagem, para ajudar o professor a identificar dúvidas dos alunos de forma automatizada. Essa abordagem auxilia tanto o professor, reduzindo o tempo empregado para responder todos os questionamentos dos alunos, quanto o aluno, indicando materiais de estudo que possam auxiliá-lo na resolução da sua dúvida.

Como trabalhos futuros, iremos: (i) desenvolver uma forma de eliminar as postagens de dúvidas que não são relacionadas com algum assunto do conteúdo programático da disciplina; (ii) realizar experimentos de extração de assuntos, utilizando outras fontes além do Wikipédia para coletar os termos relacionados.

7. Referências

- Akyuz, H. I. and Kurt, M. (2010). *Effect of teacher's coaching in online discussion forums on students' perceived self-efficacy for the educational software development. Procedia – Social and Behavioral Sciences*, 9(0):633 – 637.
- Cheng, C. K., Par, D. E., Collimore, L.-M., and Joordens, S. *Assessing the effectiveness of a voluntary online discussion forum on improving students course performance. Computers & Education*, 56(1):253 – 261, 2011.
- Dillenbourg, P. and Schneider, D. P. S. (2002). *Virtual learning environments. In 3rd Conference on Information & Communication Technologies in Education*, pages 3-18.
- Frank, E., Hall, M., Holmes, G., Kirkby, R., Pfahringer, B., Witten, I. H., and Trigg, L. (2010). *Weka-a machine learning workbench for data mining. In Data Mining and Knowledge Discovery Handbook*, pages 1269–1277. Springer.
- Friedman, N., Geiger, D., and Goldszmidt, M. (1997). *Bayesian network classifiers. Machine learning*, 29(2-3):131–163.
- Gerosa, M. A., Filippo, D., Pimentel, M., Fuks, H., and Lucena, C. J. (2010). *Is the unfolding of the group discussion off-pattern? improving coordination support in educational forums using mobile devices. Computers & Education*, 54(2):528 – 544.
- Hotho, A., Nurnberger, A., and Paas, G. (2005). *A brief survey of text mining. Journal for Computational Linguistics and Language Technology*, 20(1):19–62.
- Kim, J. *Influence of group size on students' participation in online discussion forums. Computers & Education*, 62(0):123 – 129, 2013.
- Lin, F.-R., Hsieh, L.-S., and Chuang, F.-T. (2009). *Discovering genres of online discussion threads via text mining. Computers & Education*, 52(2):481–495.
- Mohamad, S. K. and Tasir, Z. (2013). *Educational data mining: A review. Procedia - Social and Behavioral Sciences*, 97(0):320 – 324. *The 9th International Conference on Cognitive Science*.
- Oliveira Júnior, R. L., Esmín, A. A., Coelho, T. A., Araújo, D. L., Silva, L. A., Giroto, R. Uma Ferramenta de Monitoramento Automático de Mensagens de Fóruns em Ambientes Virtuais de Aprendizagem. Anais do XXII SBIE – XV WIIE, 2011.
- Ravi, S. e Kim, J. *Profiling Student Interactions in Threaded Discussions with Speech Act Classifiers. na: AI IN EDUCATION CONFERENCE (AIED)*, 2007.
- Rolim, V. B., Ferreira, R., Cordeiro, F. R. (2014). Reconhecimento de Padrões Aplicados a Comentários de Fóruns Educacionais. Encontro nacional da inteligência artificial e computacional.
- Salton, G. e Buckley, C. *Term-weighting approaches in automatic text retrieval. Journal information processing and management: an international Journal*, vol. 24, cap. 5, 1988, pp. 513-523.