

Avaliar para Nivelar e Formar : Um Sistema *Online* de Avaliação Formativa para Alunos de Biblioteconomia

Márcia G. de Oliveira¹, Elias Oliveira²

¹ Programa de Pós Graduação em Informática (PPGI)
Universidade Federal do Espírito Santo (UFES)
Caixa Postal 5011 – 29.060-970 – Vitória – ES – Brazil

²Departamento de Ciências da Informação – UFES
Vitória – ES – Brazil

clickmarcia@gmail.com, elias@lcad.inf.ufes.br

Abstract. *A major problem of teaching-learning of documents classification in Library Studies' taught modules has been how to handle increasingly large and unbalanced group of students accordingly to their levels of learning. There has already been some literature work on building student's profile. This research aims at going a step further when rather than just reveal the unbalance difficulties of learning among the pupils, proposes the reduction of them by acting quickly on their causes. To do so, we developed a Formative Assessment System for diagnosis, individualization and regulation of the teaching-learning process so that the difficulties in class can be early dealt with and the success of learning can also be promoted. The first results already pointed out the reduction of disequilibrium between groups of students and great progress of learning.*

Keywords: *Formative Assessment, Diagnosis, Levelling, Clustering, Assessment System*

Resumo. *Um grande problema do ensino-aprendizagem de classificação de documentos em cursos de Biblioteconomia tem sido lidar com turmas cada vez maiores e mais desiguais em relação aos níveis de aprendizagem. Embora já existam projetos de Educação a Distância que possibilitam reconhecer perfis de alunos e diferenciar o ensino, este trabalho, mais do que revelar as desigualdades, propõe reduzi-las. Para isso, foi desenvolvido um Sistema Online de Avaliação Formativa para diagnóstico, individualização e regulação do processo de ensino-aprendizagem com o objetivo de nivelar uma turma e promover êxitos de aprendizagem. Os primeiros resultados já apontaram para redução das desigualdades entre alunos e progressos de aprendizagem.*

Palavras-Chave: *Avaliação Formativa, Diagnóstico, Nivelamento, Clusterização, Sistema de Avaliação*

1. Introdução

As pesquisas mais recentes na área de Educação a distância têm revelado ambientes de aprendizagem cada vez mais interativos, cooperativos [Nevado et al. 2007] e adaptáveis às necessidades individuais de aprendizes [Abell 2006, Pimentel et al. 2003]. Há, desse

modo, um grande interesse em agregar aos ambientes de aprendizagem metodologias e tecnologias inteligentes essencialmente focadas no ensino e na aprendizagem. No entanto, a avaliação, sendo um método de adquirir e processar evidências necessárias para melhorar o ensino e a aprendizagem [Bloom et al. 1975], ainda pouco tem sido contemplada por essas inovações tecnológicas.

As metodologias de avaliação para ambientes de aprendizagem propostas na literatura têm focado a avaliação formativa para identificação de perfis de alunos, ensino diferenciado [Pimentel et al. 2007] e acompanhamento de aprendizagens [Reichert et al. 2008]. Essas metodologias, embora identifiquem as desigualdades entre aprendizes e proponham ambientes de aprendizagem adaptáveis às diferenças individuais, não atentam de forma expressiva para a redução dessas desigualdades.

Deve-se ainda considerar que adaptar um ambiente de aprendizagem aos diferentes níveis de conhecimento e ritmos de aprendizagem [Pimentel et al. 2007], pode, por um lado, favorecer a aprendizagem individual. Mas, por outro lado, pode impactar de forma negativa sobre a construção de aprendizagens cooperativas [Nevado et al. 2007], uma vez que aprendizes estarão em patamares desiguais de aprendizagem.

Em domínios de conhecimento como a classificação de documentos, por exemplo, reduzir as desigualdades de níveis de aprendizagem torna-se uma necessidade. Isso porque a construção de conhecimentos é favorecida pela cooperação no processo de aprendizagem [Nair 2006]. Dessa forma, não atentar para essas desigualdades nessa área de conhecimento pode comprometer a formação do profissional de classificação.

Uma proposta para resolver esse problema é um sistema computacional de apoio ao ensino-aprendizagem de classificação de documentos que possibilite nivelar uma turma e promover êxitos coletivos de aprendizagem. Para isso, propõe-se, através das avaliações diagnóstica e formativa apoiadas por tecnologia, diagnosticar o ensino-aprendizagem, individualizar o ensino e regular a aprendizagem [Perrenoud 1999].

O cenário para a primeira aplicação da nossa metodologia foi uma turma da disciplina *Representação Temática III* do curso de Biblioteconomia da Universidade Federal do Espírito Santo. Nessa disciplina estudam-se técnicas de classificação de documentos. Desse modo, através de um Sistema *Online* de Atividades de Classificação, o *SOAC*, foram oferecidas a essa turma, para cada técnica de classificação em estudo, vários exercícios de classificação. Esses exercícios foram propostos em duas etapas de avaliação: diagnóstico e avaliação formativa.

Os primeiros resultados de ambas as etapas foram submetidos a um algoritmo de clusterização automática para formação de pequenos grupos de alunos com características semelhantes e para reconhecimento de dificuldades desses grupos. Os resultados da clusterização automática da primeira etapa apontaram para grandes lacunas de aprendizagem e grupos de aprendizes explicitamente desiguais. Já para a segunda etapa, os resultados revelaram menos lacunas de aprendizagem e grupos de aprendizes menos distintos entre si. Com esses resultados, demonstramos a eficácia da Avaliação Formativa para promover êxitos de aprendizagem e para reduzir desigualdades de níveis de aprendizagem.

As seções deste trabalho foram organizadas conforme a ordem a seguir. Na Seção 2, descrevemos os problemas do ensino-aprendizagem de classificação em cursos de Biblioteconomia. Na Seção 3, apresentamos a nossa proposta metodológica de avaliações

diagnóstica e formativa apoiadas por tecnologia. Na Seção 4, é apresentado um estudo real da aplicação da nossa metodologia em uma turma de Biblioteconomia. Na Seção 5, são relatados os experimentos e os resultados alcançados antes e depois da avaliação formativa. Na Seção 6, apresentam-se as considerações finais.

2. A Realidade do Ensino-Aprendizagem de Classificação em Biblioteconomia

O ensino-aprendizagem de classificação em cursos presenciais de Biblioteconomia tem sido um grande desafio. Além da formação a cada ano de turmas cada vez maiores, os alunos ingressantes trazem consigo bases diferenciadas de conhecimentos que evidentemente influenciam no rendimento desses alunos ao longo de um curso de classificação.

Essas desigualdades no ensino-aprendizagem de classificação impactam diretamente na formação do classificador pois, de acordo com [Nair 2006], o conhecimento do classificador se molda com a experiência e com o conhecimento gerado por outros classificadores em processo cooperativo de aprendizagem.

Além das desigualdades, dificulta também o ensino-aprendizagem de classificação o modelo pedagógico ainda baseado em práticas behavioristas [Aquino 2008] e sem auxílio computacional. Desse modo, um professor acaba por não ter condições de acompanhar a aprendizagem de um grande número de alunos com necessidades individuais tão diferenciadas e muito menos de promover a aprendizagem cooperativa.

3. Uma Avaliação a favor da Aprendizagem e contra as Desigualdades

Diante dos problemas apresentados no ensino-aprendizagem de classificação de documentos, torna-se inadequado, em especial na área de classificação da informação, manter um modelo pedagógico que preserve as desigualdades de níveis de aprendizagem.

Como nos aponta [Ballester 2003], se quisermos mudar essa prática pedagógica, é necessário mudar a prática da avaliação, ou seja, mudar sua finalidade, o que e como se avalia. Para isso, a avaliação, como ação pedagógica, deve cumprir três funções didático-pedagógicas [Ballester 2003]: Diagnóstica, Formativa e Somativa.

A avaliação diagnóstica deve ter o papel de identificar habilidades e dificuldades de aprendizagem além de reconhecer perfis de alunos. A avaliação formativa, por sua vez, deve consistir de *feedbacks* e ajustes no processo de ensino-aprendizagem para alcançar objetivos traçados. Já a avaliação somativa tem como função classificar aprendizes de acordo com níveis de aproveitamento no processo de aprendizagem.

Na Figura 1 apresentamos o modelo avaliativo [Perrenoud 1999, Ballester 2003] em que se baseia a nossa proposta metodológica para atender às duas funções didático-pedagógicas da avaliação. Através dessa proposta visamos alcançar, com o apoio da tecnologia, as áreas escurecidas em destaque na Figura 1, áreas essas em geral negligenciadas nos atuais processos avaliativos.

Para o desenvolvimento de um Sistema *Online* de Avaliação que atendesse à nossa proposta metodológica de avaliação, levamos em consideração que o principal instrumento de toda avaliação formativa é e continuará sendo o professor comprometido em uma iteração com o aluno [Perrenoud 1999]. Consideramos ainda que a pedagogia diferenciada e a avaliação formativa que a acompanha, conforme afirma Phillippe Perrenoud

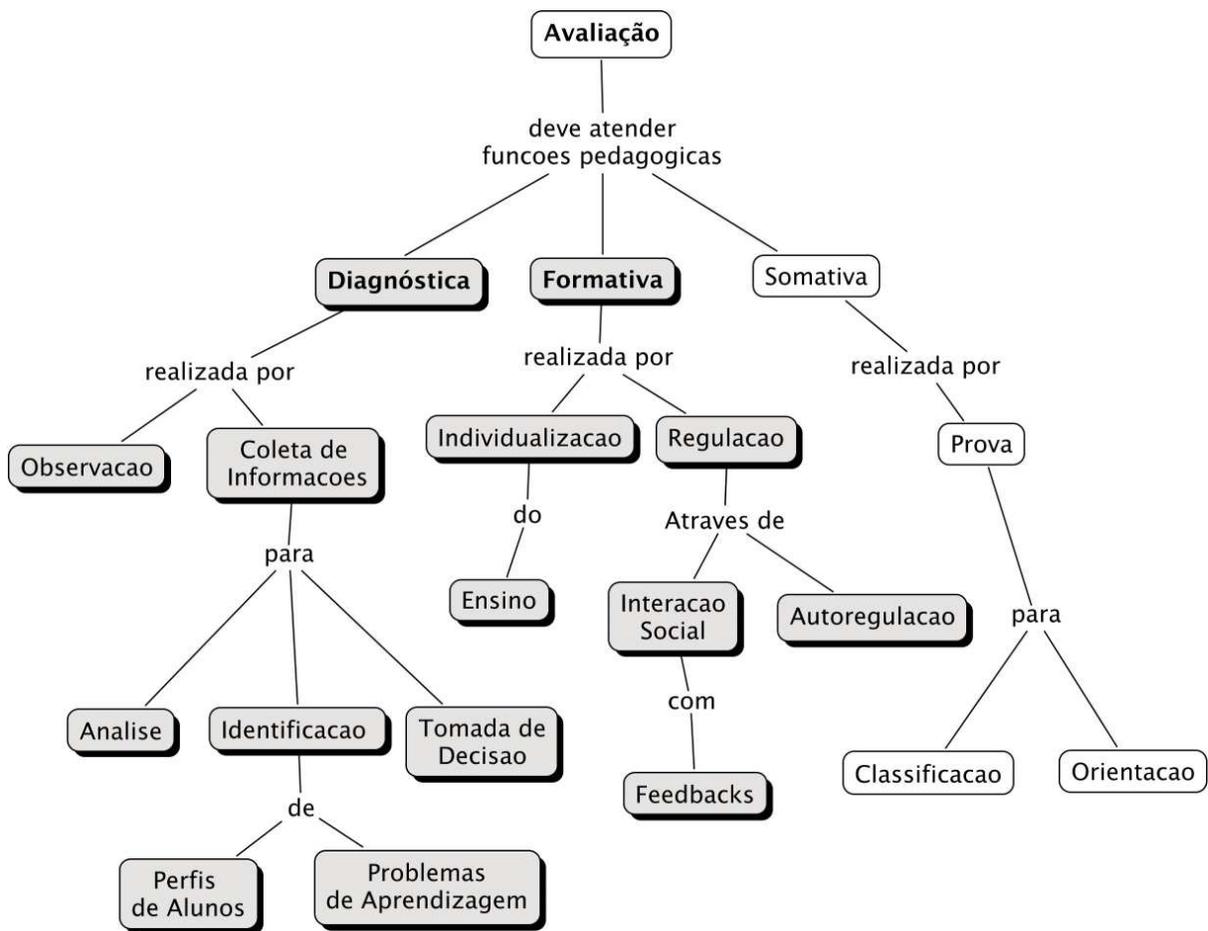


Figura 1. Modelo de Avaliação

[Soares and Velloso 2007], exigem competências de observação, de metacognição e didáticas direcionadas a propor tarefas e situações pertinentes e produtivas.

Dessa forma, entendemos que diferenciar o ensino [Perrenoud 1999] não é adaptá-lo às desigualdades de níveis e ritmos de aprendizagem, mas sim mudar a prática pedagógica. Essa mudança consiste em individualizar o ensino através de constantes *feedbacks* de iterações entre professor e aluno com a finalidade de regular a aprendizagem.

4. Estudo de Caso

Para estudos da aplicação das avaliações diagnóstica e formativa participaram os alunos e o professor da disciplina *Representação Temática III* do Departamento de Ciências da Informação da Universidade Federal do Espírito Santo. Nessa disciplina os alunos aprendem as técnicas de classificação de documentos como, por exemplo, a *CDD* (Codificação Decimal de Dewey) e a *CDU* (Codificação Decimal Universal).

O Sistema *Online* de Atividades de Classificação, o *SOAC*, foi utilizado pelo professor e alunos da turma em estudo durante o primeiro semestre do ano de 2008 como ferramenta de apoio ao ensino-aprendizagem por avaliação formativa.

A Figura 2 é uma representação em mapa conceitual do modelo de Sistema *Online* de Avaliação Diagnóstica e Formativa *SOAC* para atender a nossa proposta pedagógica.

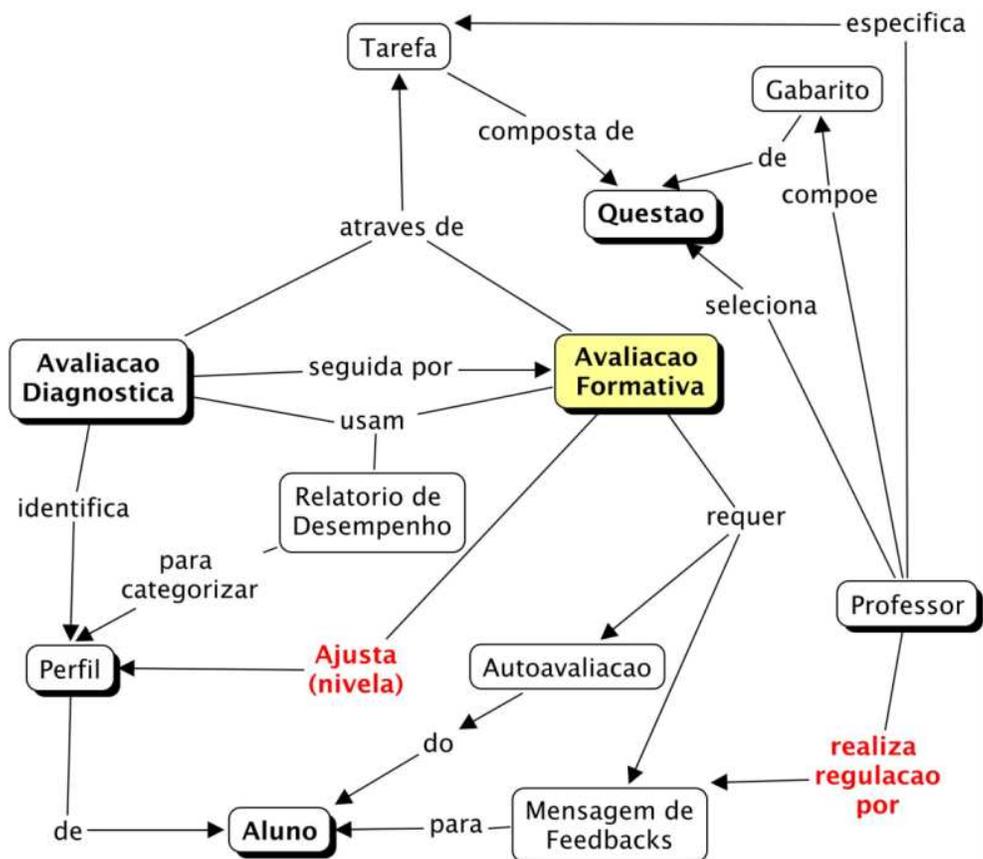


Figura 2. Modelo de Sistema Online de Avaliação Formativa

Para as primeiras atividades de classificação, foram utilizadas notícias de jornal para os alunos atribuírem classes. Para isso, cerca de 2700 notícias do Jornal *A Tribuna* foram selecionadas e disponibilizadas no SOAC. Essas notícias estavam pré-classificadas em nove classes: *Cidades*, *Especial* (notícias especiais), *Política*, *Economia*, *Televisão*, *Família*, *AT2* (Caderno 2), *Informática* e *Miscelânea* (outros assuntos). No Sistema SOAC, essas classes foram representadas, respectivamente, por números de 1 a 9.

Após escolher os objetos de classificação, realizar um treinamento para uso do Sistema SOAC e especificar a tarefa de classificação de notícias, o professor liberou a tarefa durante duas semanas para os alunos realizarem as atividades de classificação.

5. Experimentos e Resultados

Depois de os alunos da turma de *Representação Temática III* em estudo concluírem a primeira etapa de atividades de classificação de notícias, os resultados de desempenhos de cada aluno foram computados pelo sistema para serem analisados pelo professor.

Para a fase de individualização da avaliação formativa, foram enviadas mensagens eletrônicas para cada aluno informando os seus desempenhos na tarefa e as suas falhas. Para a fase de regulação, foram apresentados à turma os critérios de classificação de notícias do Jornal *A Tribuna* e por que algumas dessas classes são mais difíceis.

Após realização das fases de avaliação formativa, a tarefa de classificar notícias foi novamente liberada no sistema SOAC (Figura 3).



Figura 3. Sistema Online de Atividades de Classificação

5.1. Clusterização Automática

A clusterização utiliza a aprendizagem não-supervisionada para realizar o agrupamento, isto é, a auto-organização de padrões em classes não-rotuladas (*clusters*) [Jain et al. 1999].

O *K-means* é uma das técnicas mais aplicadas na área de classificação de padrões. Para este trabalho, no entanto, optamos por aplicar a técnica *Bisecting K-means*, uma variação do *K-means*, por formar *clusters* de tamanhos mais uniformes. O algoritmo *Bisecting K-means* funciona da seguinte forma [Steinbach et al. 2000]:

1. Dividir um *cluster* em dois *sub-clusters* utilizando o *K-means* [Looks et al. 2007]
2. Repetir o passo 1 por I vezes e escolher a melhor divisão por similaridade.
3. Repetir os passos 1 e 2 até alcançar o número de *clusters* desejado.

A complexidade de tempo desse algoritmo, utilizando o cálculo de cosseno como medida de similaridade entre padrões, é $O(NNZ * \log(k))$, onde NNZ é o número de não-zeros da matriz de entrada e k é o número de *clusters* formados.

5.2. O Antes e o Depois da Avaliação Formativa

As planilhas com os resultados das avaliações diagnóstica (ver Figura 4) e formativa foram submetidas a algoritmos de clusterização [Steinbach et al. 2000] do *Software Cluto*¹ para obter relatórios dos níveis de aprendizagem dos alunos antes e depois da avaliação formativa.

¹Software e documentação disponíveis em : <http://glaros.dtc.umn.edu/gkhome/views/cluto>

Resultados de Desempenho dos alunos em Atividades de Classificação de Notícias - Diagnóstico									
	C1	C2	C3	C4	C5	C6	C7	C8	C9
João	0.5	0	0	1	0	0	0	0.33	0.40
Pedro	0.33	1	0.46	1	0	1	0	0.42	0
André	0.33	0.2	0.11	0.33	0.2	0.92	0	0.34	0.83
Luiz	0	0	0	0	0	0	0	0.7	1

Figura 4. Desempenho de Alguns Alunos em Atividades de Classificação: Diagnóstico

A planilha obtida (Figura 4) apresenta os índices de desempenhos dos alunos que variam de 0 a 1 (ou de 0 a 100%), em cada uma das nove classes de notícias C_i (para $i = 1, 2, \dots, 9$) na etapa de diagnóstico.

Os resultados da clusterização fornecidos pelo *Software Cluto* para as avaliações diagnóstica e formativa apresentaram a turma em estudo dividida em cinco grupos (*clusters*). Escolhemos arbitrariamente formar cinco grupos de alunos por motivos didáticos.

Esses grupos foram organizados por níveis similares de desempenhos dos aprendizes nas atividades de classificação. Os gráficos da Figura 5 apresentam, respectivamente, os resultados das avaliações diagnóstica (o *Antes*) e formativa (o *Depois*).

Nesses gráficos, os alunos são representados por linhas e as classes de notícias por colunas. Os níveis de desempenhos dos alunos nas atividades de cada classe C_i são representados pelas tonalidades de uma cor. Desse modo, quanto mais intensa for a tonalidade dessa cor nas classes C_i , melhor desempenho um aluno obteve nas questões dessas classes. Da mesma forma, quão mais próximas forem essas tonalidades da cor branca, menores os índices de desempenhos de um aluno nessas questões.

No primeiro gráfico da Figura 5 há grandes concentrações de áreas brancas indicando lacunas de aprendizagem. Além disso, observa-se uma clara distinção entre os cinco grupos formados, o que revela desigualdades de níveis de aprendizagem.

Analisando as áreas brancas do primeiro gráfico da Figura 5, observamos que elas se concentram predominantemente nas classes *Miscelânea* (Classe C2), *Especial* (Classe C3) e *Família* (Classe C7). Pelas características particulares dessas classes, os insucessos de aprendizagem nelas identificados sugerem dificuldades dos aprendizes na etapa de indexação, a etapa que antecede a fase de classificação de documentos.

A indexação é definida como o ato de descrever ou identificar um documento em termos de seu conteúdo. Para indexar documentos, são necessárias a leitura documentária e a análise do assunto para a extração de termos significativos em textos [Fujita 1999].

As deficiências em reconhecer a classe C2 podem ter origem na dificuldade de análise do assunto pois essa classe é representada por assuntos que não se enquadram nas outras classes. A classe C7 pode ser também um indicador de deficiências de análise de assunto pois os termos que representam as notícias dessa classe, em geral, estão também

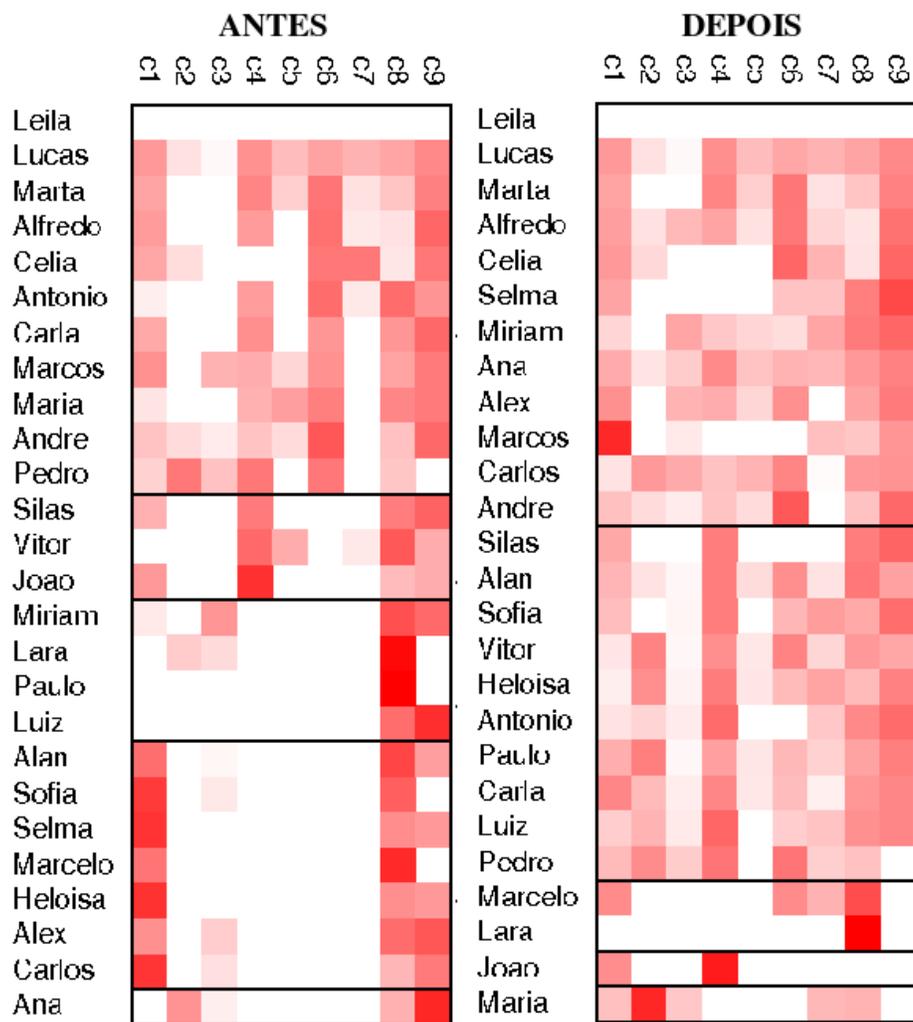


Figura 5. O Antes e o Depois da Avaliação Formativa

associados a outras classes. Desse modo, exige-se que o classificador analise esses termos em conjunto e verifique se eles se aplicam a um contexto de família.

Já as grandes áreas brancas na classe C3 podem estar revelando deficiências na leitura documentária e, por conseguinte, na análise de assunto e na extração de termos representativos. Isso porque as notícias da classe C3 são grandes e seus termos podem representar qualquer classe. Além disso, é um indicador de dificuldade de análise de assunto não reconhecer que as notícias da classe C3 são "especiais".

Após o diagnóstico das deficiências e a aplicação da avaliação formativa, de acordo com o segundo gráfico da Figura 5, as áreas brancas, antes em grandes concentrações, reduziram-se consideravelmente. Por conseguinte, os cinco novos grupos formados tornaram-se menos distintos entre si. Observamos ainda, que a maior parte dos alunos que antes estavam nos grupos com as maiores concentrações de áreas brancas, no depois da avaliação formativa melhoraram os seus desempenhos.

A Tabela 1 apresenta os resultados de clusterização automática antes e depois da avaliação formativa. Os grupos (*clusters*) são numerados de 0 a 4 e visualizados na Figura

Avaliação Formativa

ANTES			DEPOIS		
Cluster	Tamanho	ISIM	Cluster	Tamanho	ISIM
0	1	100%	0	1	100%
1	7	89.1%	1	1	100%
2	4	79.9%	2	2	85.1%
3	3	88.4%	3	10	86.3%
4	11	68.4%	4	12	68.6%

Tabela 1. Resultados de Clusterização Automática Antes e Depois da Avaliação Formativa

5 de baixo para cima nos dois gráficos. O valor *ISIM* indica o índice de similaridade interna entre os elementos de um mesmo grupo. No primeiro gráfico da Figura 5, no *Antes* da avaliação formativa, observamos que o grupo de 11 alunos (*cluster* 4) que apresentaram os melhores desempenhos possui 68.40% de *ISIM* no *Antes* da Tabela 1 e a aluna Leila não obteve desempenho. Ainda na Figura 5, no *Depois*, visualizamos êxitos de aprendizagem da maioria dos alunos e redução de desigualdades entre eles. Podemos constatar isso através dos tamanhos e dos valores de *ISIM* dos *clusters* 3 e 4 no *Depois* da Tabela 1. Eles revelam um número maior de alunos mais semelhantes entre si nos grupos com melhores desempenhos. Além disso, os tamanhos dos *clusters* 1 e 2 no *Depois* da Tabela 1 indicam reduções dos grupos de alunos com baixos desempenhos. Isso nos leva a concluir que, de fato, houve redução de desigualdades e êxitos coletivos de aprendizagem foram alcançados após a avaliação formativa.

6. Considerações Finais

Ainda que a ação formativa seja um processo difícil de se gerenciar, considerando turmas com um grande número de alunos e com níveis de aprendizagem desiguais, ela é possível e menos complexa se apoiada pela tecnologia, conforme apresentamos neste trabalho.

Os resultados apresentados da avaliação formativa demonstram a eficácia da metodologia proposta neste trabalho. Antes, ao longo de um semestre, um aluno não resolvia mais que 30 questões de classificação. Agora, com o apoio da nossa tecnologia, foram resolvidos e, automaticamente corrigidos, cerca de 150 exercícios por aluno. Antes da avaliação formativa, havia grupos de alunos mais distintos entre si e lacunas de aprendizagem. Depois, os grupos de alunos tornaram-se menos distintos indicando nivelamento, e as lacunas de aprendizagem reduziram-se.

Neste trabalho a avaliação diagnóstica revelou deficiências de aprendizagem tendo como variáveis de avaliação apenas classes de notícias. Embora as características de algumas dessas classes tenham apontado para deficiências cognitivas, essencialmente diagnosticamos dificuldades de aquisição de conhecimentos. Como trabalho futuro pretendemos utilizar variáveis que possam revelar e medir dificuldades de domínio cognitivo.

Uma vez que este trabalho possibilitou a identificação de problemas de aprendizagem, a individualização do ensino e a regulação de aprendizagens [Perrenoud 1999], agregamos qualidade ao ensino-aprendizagem de classificação. Esperamos que este trabalho seja uma contribuição para a redução das desigualdades que dificultam o processo de aprendizagem e para a formação de competências nessa área do conhecimento.

Referências

- Abell, M. (2006). Individualizing Learning Using Intelligent Technology and Universally Designed Curriculum. *Journal of Technology, Learning and Assessment*, 5(3):7–10. Lynch School of Education, Boston College.
- Aquino, M. A. (2008). As Contribuições da Educação aos Processos Formativos na Ciência da Informação. *Transinformação Jan/Abr*, 20(1):59–71.
- Ballester, M. (2003). *Avaliação como Apoio à Aprendizagem*. Artmed. Trad. Valério Campos. Porto Alegre, RS.
- Bloom, B., Hastings, J., and Madaus, G. (1975). *Evaluación del Aprendizaje*. Buenos Aires, Troquel.
- Fujita, M. (1999). A leitura do indexador: Estudo de observação. volume 4, pages 101–116. *Perspectivas em Ciência da Informação*.
- Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering: a review. *ACM Comput. Surv.*, 31(3):264–323.
- Looks, M., Levine, A., Covington, G., Loui, R., Lockwood, J., and Cho, Y. (3-10 March 2007). Streaming Hierarchical Clustering for Concept Mining. *Aerospace Conference, 2007 IEEE*, pages 1–12.
- Nair, M. (2006). A Organização da Informação e do Conhecimento a partir do Conhecimento Explícito de Classificadores. Monografia de Especialização em Gestão Estratégica do Conhecimento e da Inovação. Departamento de Ciências da Informação, Universidade Federal do Espírito Santo(UFES), ES.
- Nevado, R., Carvalho, M., and Menezes, C. (2007). *Aprendizagem em Rede na Formação à Distância*. Porto Alegre: Ricardo Lenz.
- Perrenoud, P. (1999). *Avaliação: Da Excelência à Regulação das Aprendizagens – Entre Duas Lógicas*. Artmed Editora, Porto Alegre, RS.
- Pimentel, E., Alves, A., Oliveira, B., Ikebara, D., Bottaro, P., and Lopes, R. (2007). Avaliações adaptativas baseadas no nível de aquisição de conhecimentos do aprendiz. In *Anais do XVIII Simpósio Brasileiro de Informática na Educação*, pages 566–574. Sociedade Brasileira de Computação(SBC).
- Pimentel, E., França, V., and Omar, N. (2003). A Caminho de um Ambiente de Avaliação e Acompanhamento Contínuo de Aprendizagem em Programação de Computadores. In *II Workshop de Educação em Computação e Informática do Estado de Minas Gerais (WEIMIG 2003)*, pages 212–213, Poços de Caldas, MG.
- Reichert, C., Schirmbeck, F., and Lemos, M. (2008). Acompanhamento e Avaliação da Aprendizagem de Alunos em Cursos EAD: uma Ferramenta de Gestão. In *Anais do V Congresso Brasileiro de Ensino Superior a Distância*. UFRGS/FAURGS, RS.
- Soares, A. and Velloso, A. (2007). Entrevista: Phillippe Perrenoud. volume 2, page 6. *Revista EmFormação*.
- Steinbach, M., Karypis, G., and Kumar, V. (2000). A Comparison of Document Clustering Techniques. In *KDD Workshop on Text Mining*.