

Analizando Diálogos para Classificação de Padrões Utilizando Redes Neurais Artificiais e Árvores de Decisão

Ana Cláudia H. Vieira¹, Patrícia C. Tedesco¹, Aline Timóteo¹, Allan Lima¹

¹Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7851 – 50732-970 – Recife – PE – Brasil

{achv, pcart, alt, adsl}@cin.ufpe.br

Resumo. A maioria dos sistemas de educação a distância apenas oferecem informações e repassam conhecimento aos alunos. Eles não fornecem mecanismos de monitoramento para que os usuários se interessem em utilizá-los. Foi criada, então, uma ferramenta que monitora o diálogo dos usuários, analisando as características particulares das suas participações, e classifica o ponto de vista de aprendizagem. Para classificação, foi utilizada uma arquitetura complementar de Redes Neurais Artificiais e Árvores de Decisão. Os resultados práticos alcançados mostraram ser possível identificar os assuntos que despertam maior interesse dos alunos e os que causam mais dúvidas, ajudando a melhorar o processo de ensino/aprendizagem.

Palavras-chave: classificação de diálogos, aprendizagem.

Abstract. Most distance educational systems only supply students with information and knowledge. They do not provide monitoring ways to make users interested in using them. So, It was developed a tool that monitors users dialogues, analyzing particular features of each person, and then classify the dialogue according to Its learning degree. As classifiers, It was used a complementary architecture of Artificial Neural Networks and Decision Trees. Practical results showed a true possibility of identifying the most interesting subjects for student and those harder to be understood. This helps the teaching/learning process.

Keywords: dialogue classification, learning.

1. Introdução

Com a evolução dos meios de comunicação em especial a Internet, surgiram idéias de como melhor aproveitar os recursos tecnológicos para o crescimento humano. Uma das maneiras muito divulgadas hoje em dia é o ensino à distância, que permite que os usuários aprendam e ensinem, sem a necessidade de estarem fisicamente unidos.

Um aspecto interessante do compartilhamento de conhecimento são os *chats* educacionais, diálogos que permitem a comunicação de maneira síncrona entre professores e alunos, permitindo a troca de idéias, opiniões e agrupamento de interesses comuns. Na maioria das vezes os *chats* são utilizados apenas como ferramentas de “bate-papo” sem nenhuma preocupação com a aprendizagem.

De acordo com as habilidades presentes na fala [Searle 1972] pode-se descobrir diversos aspectos relacionados a efetividade dos diálogos, o que nos permite realizar

uma análise do quanto o diálogo foi produtivo para o fim a que foi destinado ou o quanto não foi produtivo.

Uma das formas de análise destes diálogos pode ser feita de acordo com o padrão que os diálogos apresentam. Neste trabalho são apresentadas as Redes Neurais Artificiais [Haykin 1994] e as Árvores de Decisão [Russel e Norvig 2003] para classificar diálogos de acordo com o aprendizado que estes proporcionam. Segundo Boud e colegas [1985], a reflexão é uma das habilidades mais importantes para a aprendizagem. Enquanto se está refletindo sobre determinado assunto, normalmente se está aprendendo. Para verificar se houve ou não reflexão durante o diálogo, são analisados alguns aspectos existentes nas conversas. Assim, para este trabalho, um diálogo é efetivo quando trata do assunto proposto a ser discutido e apresenta indicativos de reflexão.

Este artigo está organizado da seguinte forma: na seção 2, é apresentada a ferramenta de aprendizagem; na seção 3 são mostradas as abordagens de classificação utilizadas (Redes Neurais Artificiais e Árvores de Decisão); na seção 4 está a discussão dos resultados dos testes realizados; a seção 5 é composta das conclusões e trabalhos futuros.

2. Ferramenta de Aprendizagem

Um dos principais problemas encontrados atualmente nos ambientes de educação a distância é a falta de motivação dos alunos. O professor muitas vezes não consegue transmitir virtualmente o que conseguiria em salas de aula presenciais, pois virtualmente perdem-se diversos canais de comunicação, tais como os sons da fala, as expressões corporais, a visualização dos gestos de cada pessoa. Atualmente existem funcionalidades que são aplicadas aos ambientes de educação a distância que podem auxiliar na comunicação, tais como, indicações de sentença, emoções (também conhecidas como *smiles*), a verificação das habilidades de fala presentes nos diálogos e o controle de tempo de utilização.

Sabendo-se que a maioria dos alunos quando se sentem motivados aprendem melhor, criou-se uma ferramenta de auxílio ao aprendizado, que indica para o professor como os alunos estão se comunicando. Uma das formas para a verificação da existência de aprendizado através dos diálogos é a constatação da discussão do assunto proposto e da obtenção de reflexão por parte dos usuários.

Uma das formas mais comuns de comunicação através da Internet são os *chats*, ou salas de “bate-papo”. Os ambientes educacionais utilizam-se deste recurso para que os alunos possam discutir assuntos desenvolvidos em sala de aula ou desenvolver trabalhos em grupo, promover debates, entre outros. Através das conversas desenvolvidas nos *chats*, pode-se dar ao professor e aos alunos um indicativo de como o diálogo decorreu em termos de reflexão. Assim, o professor pode direcionar melhor sua prática de ensino, já que através das habilidades de fala presentes no diálogo, é possível identificar quais são os assuntos de maior interesse dos alunos e em quais encontram maiores dificuldades. Além disto, identificar quais alunos são mais comunicativos, os que possuem maiores dificuldades em se expressar, aqueles que somente concordam e os que discordam de tudo.

O monitoramento das conversas não é realizado na maioria dos *chats*, muitas vezes são utilizados agentes, tutores virtuais e sistemas unidos aos *chats* para realizar

este monitoramento. Para facilitar a utilização dos recursos de monitoramento dos diálogos, foi criada uma ferramenta de auxílio ao aprendizado, a Oxentchê-chat.

A Oxentchê-chat [Vieira 2004] permite que o professor receba um *feedback* das conversas ocorridas durante a utilização do sistema, assim como, que cada aluno tenha idéia de como foi sua participação. A ferramenta possui um módulo de *feedback on-line*, que pode ser recebido pelos professores durante a ocorrência do diálogo e um *feedback off-line*, um módulo que permite que o professor verifique a resposta do sistema após o período de utilização *on-line* da ferramenta.

2.1. Desenvolvimento da Oxentchê-chat

Inicialmente foram coletados diversos diálogos disponíveis na Internet para formar uma base de diálogos. Todos os diálogos tinham propósitos educacionais, ou seja, todos foram obtidos de sites, salas de bate-papo, disciplinas ou cursos a distância que tinham como foco o aprendizado de assuntos da área de informática.

Após a coleta, todos os diálogos foram lidos e anotadas informações destes, tais como, o assunto conversado, os marcadores conversacionais presentes [Marcuschi 2003], tamanho do diálogo, número de participantes. Assim, foi possível dar início aos testes para a implementação da ferramenta.

A ferramenta foi implementada em Java e XML, desenvolvida em módulos, facilitando assim acrescentar mais funcionalidades futuramente. O assunto proposto é verificado por uma ontologia que possui 6 domínios (Introdução a Inteligência Artificial, Agentes Inteligentes, Aprendizado de Máquina, Representação do Conhecimento, Sistemas Multi-agentes e interface). Através do *log* do diálogo, o sistema faz uma comparação com as palavras disponíveis na ontologia do domínio especificado pelo professor, que deve corresponder ao assunto proposto para o diálogo. Existe um sistema de busca em profundidade para verificar o nível e a quantidade de vezes que as palavras aparecem na conversa, pois algumas palavras do domínio podem estar presentes no diálogo, mas não indicarem que o assunto discutido foi o mesmo proposto.

Após a verificação do assunto proposto são contados os participantes, o número de intervenções e as habilidades de fala, que são identificadas através dos marcadores (do inglês *sentence openers*) que são as sentenças existentes no início das frases. Com a identificação dos marcadores é possível dizer quais habilidades de fala foram desenvolvidas em determinada conversa. Em Vieira e colegas [2004] é possível observar os detalhes da arquitetura do sistema e suas principais funcionalidades.

A seguir existem os módulos das Redes Neurais Artificiais e das Árvores de Decisão que de posse das informações obtidas anteriormente, fazem uma classificação do diálogo, enviando como resposta um *feedback* com explicações da qualidade do diálogo. O *feedback* é enviado tanto para o professor, quanto para os alunos.

O professor pode identificar quais foram os temas que mais geraram dúvidas em seus alunos, quais foram mais interessantes e quais pontos devem ser melhorados para aumentar a motivação e com isto auxiliar no aprendizado. Os alunos, por sua vez, podem verificar como foi sua participação no diálogo perante os demais colegas, recebendo informações sobre a forma que podem melhorar a conversação, auxiliando assim o processo de colaboração.

É importante ressaltar que a base de diálogos possuía inicialmente 200 diálogos, e passou a ter 400 para os testes descritos neste artigo, já que após a utilização do sistema em sala de aula, foi verificada a necessidade de novos testes para um maior detalhamento do *feedback*, assim, todo o trabalho realizado para os 200 primeiros diálogos, foi repetido para os novos diálogos, formando uma base coerente para os diversos testes realizados.

A seguir serão apresentadas as Redes Neurais Artificiais e as árvores de Decisão para classificação de diálogos.

3. Classificadores de Diálogos Educacionais

Existem diversas formas de monitorar o aprendizado em ambientes educacionais, alguns trabalhos divulgam o monitoramento por abordagens de ambientes de aprendizado inteligentes com agentes companheiros [Torreão 2004], sociedades de agentes inteligentes [Siebra 2004], Modelos Escondidos de Markov [Soller et al 2002]. Neste trabalho, são utilizadas as Redes Neurais Artificiais e as Árvores de Decisão como classificadores de padrões de diálogos.

Atualmente a utilização de mecanismos de classificação como Redes Neurais Artificiais (RNAs), possibilitam a realização da classificação de determinados padrões, com uma excelente precisão. Porém, para uma descrição mais detalhada, é necessária a utilização de uma outra abordagem. Assim, para este projeto, além das RNAs foram estudadas as Árvores de Decisão (ADs), que permitem, a partir de sua estrutura, o estabelecimento de regras que possibilitam o sistema a indicar porque uma entrada pertence a determinada saída [Russel e Norvig 2003]. Ou seja, para a análise de diálogos, as RNAs indicam a resposta se o diálogo foi ou não efetivo e as ADs, indicam a forma como isto foi constatado.

3.1 Redes Neurais para Classificação de Diálogos Educacionais

De acordo com as habilidades de fala definidas através dos marcadores conversacionais é possível obter um padrão dos diálogos educacionais. Redes Neurais Artificiais (RNAs) são amplamente utilizadas na literatura como classificadores de padrões [Haykin 1994]. Assim, decidiu-se utilizá-las para classificar os diálogos como efetivos ou não efetivos, sabendo-se que efetivos seriam os diálogos que obtiveram reflexão do assunto proposto por parte de seus participantes.

Existem diversas formas de verificar se um diálogo gerou reflexão ou não, neste trabalho, a forma escolhida foi monitorar os marcadores utilizados, levando em consideração as habilidades de fala correspondentes a cada marcador, além do número de participantes do diálogo, discussão do assunto proposto e o tipo, ou seja, se havia ou não a presença de um professor, coordenador/moderador ou convidado.

Os dados coletados de uma base de diálogos formada para o trabalho descrito neste artigo, eram de diálogos reais, disponíveis na Internet ou desenvolvidos em ambientes de ensino a distância, enviados pelos professores responsáveis para o nosso grupo de pesquisa.

As informações relevantes obtidas nesta base serviram como dados de entrada para a rede neural treinada para este sistema. Inicialmente a base foi dividida em 50% para treinamento, 25% para validação e os 25% restantes para teste de um total de 400

diálogos existentes. Os treinamentos foram realizados, variando-se a taxa de aprendizado, o número de neurônios escondidos, além das características de entrada e os conjuntos de treinamento, validação e teste. Os dados foram normalizados para que algumas características não tivessem maior peso do que outras. Foram realizados diversos treinamentos e testes com redes MLP (Multi Layer Perceptron) e RBF (Radial Base Function), os melhores resultados dos testes obtidos são mostrados na Tabela 1, a seguir.

Tabela 1 - Resultados dos Testes com as RNAs

<i>Rede</i>	<i>Erro Efetivos</i>	<i>Erro não Efetivos</i>	<i>Erro Total em Porcentagem</i>
MLP	0.258065	0.105263	0,2
RBF	0.295082	0.236842	0,27

De posse destes resultados pode-se dizer que as RNAs são capazes de classificar diálogos com até 80% de acerto, sendo assim, resolveu-se implementá-las na ferramenta Oxentchê-chat.

3.2. Árvores de Decisão para Auxílio nas Respostas

Como dito anteriormente, as RNAs indicam apenas se o diálogo foi ou não efetivo. Após alguns testes com a Oxentchê-chat em ambientes de aprendizado reais, foi identificado que a ferramenta seria mais interessante para o aprendizado se esta indicasse além da resposta, o motivo para tal. Assim, foram estudadas as Árvores de Decisão (ADs).

A Árvore de Decisão é um método que utiliza um algoritmo recursivo para subdividir o conjunto de treinamento até que uma partição seja composta inteiramente ou predominantemente de casos pertencentes a uma mesma classe, ou seja, que algum critério de parada seja alcançado. Para decidir qual atributo será utilizado em cada subdivisão, um teste estratégico é adotado como critério de quebra [Russel e Norvig 2003].

Foram realizados diversos testes com as ADs, sendo possível identificar quais as regras possíveis para indicação das respostas a serem enviadas para os usuários da Oxentchê-chat. Foram realizados treinamentos com as árvores de decisão e de acordo com os resultados obtidos foi gerada uma árvore com as regras a serem seguidas durante a classificação dos diálogos. Esta árvore foi gerada com os dados dos 400 diálogos da base. As habilidades de comunicação envio e conflito criativo [Robertson et al. 1998] se mostraram de grande importância para a qualidade efetiva do diálogo, comprovando os estudos realizados na literatura.

A Figura 1 mostra o resultado da análise de um diálogo realizada pela ferramenta Oxentchê-chat. Do lado inferior direito da tela, o professor indica se prefere analisar a conversa individual (selecionando o nome correspondente do usuário) ou analisar o diálogo completo. São mostradas diversas informações tais como o tempo de diálogo, as quantidades de cada tipo de habilidade, o número de intervenções, além da classificação final do diálogo com sua justificativa.

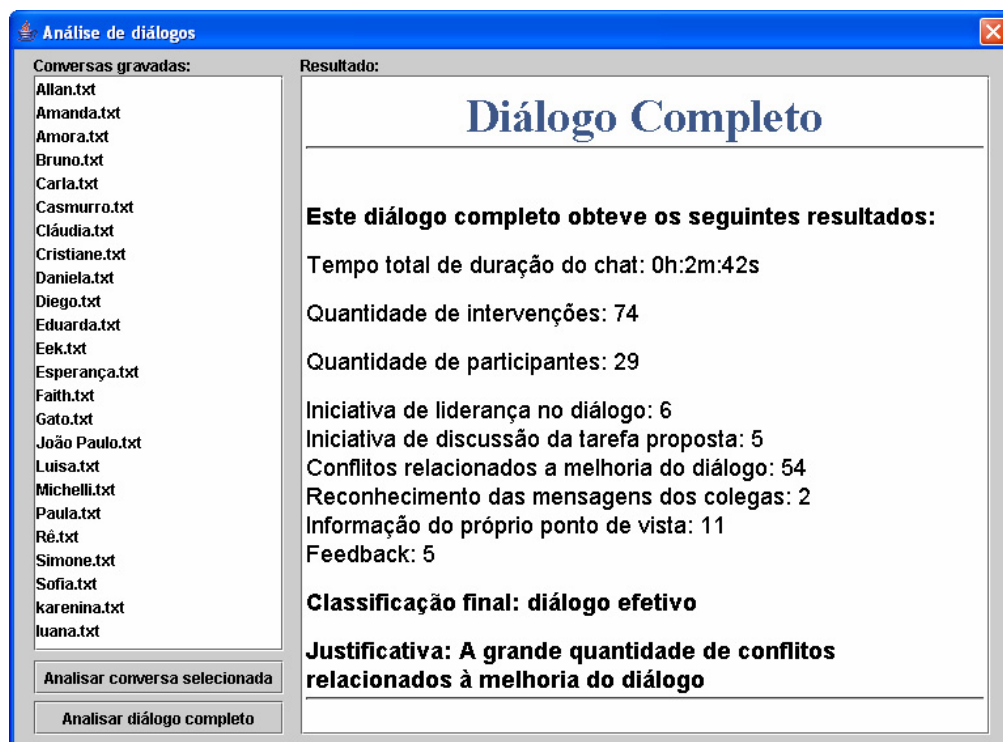


Figure 1 - Análise de um Diálogo através da Oxentché-chat

Assim, foi possível melhorar o *feedback* enviado aos professores e alunos, indicando agora o “porquê” de determinada classificação. Ainda estão sendo realizados testes com o sistema indicando o *feedback* com as respostas das RNAs unidas com as ADs, para verificar na prática se a forma de apresentação das respostas ainda pode ser melhorada.

Feedback Recebido por Alunos e Professores

Com os resultados obtidos pelas RNAs e as ADs o sistema retorna para o professor e para os alunos um *feedback* contendo a análise resultante do diálogo ocorrido.

Inicialmente o sistema apresentava respostas simples, indicando a quantidade (em números) de habilidades utilizadas e seus resultados perante o número de interações realizadas. Com a implementação das ADs foi possível enviar uma resposta mais detalhado e coerente do que o enviado somente com as funcionalidades das RNAs.

São enviados relatórios referentes a participação do usuário no diálogo, com a verificação da efetividade perante o processo de aprendizagem. O professor possui acesso ao *log* do diálogo completo e individual dos alunos. É importante ressaltar que somente o professor tem acesso aos relatórios de todos os alunos, e este gerenciamento é feito através de senhas. Os alunos têm acesso a seus relatórios individuais que possuem a indicação de como foram perante o diálogo completo, mas sem interferir nos relatórios dos colegas, por uma questão de sigilo e orientação.

4. Discussão dos Resultados

Para validação da Oxentchê-chat, foram realizados testes em sala de aula, durante as aulas, com a presença do professor e sem a presença do professor. Para verificação por parte dos professores, foram gerados o *feedback on-line* e também o *off-line*.

Os testes foram realizados na Universidade Federal de Pernambuco (UFPE) e na Universidade Federal Rural de Pernambuco (UFRPE), com alunos universitários e o monitoramento dos pesquisadores do grupo de Inteligência Artificial, da coordenadora do grupo e dos respectivos professores das turmas.

Inicialmente foram feitos testes preliminares para verificar se existiam problemas técnicos, tais como demora no processamento, problemas com o servidor, ou com os módulos clientes, entre outros. Após, foram realizados testes de usabilidade do chat, para verificação inclusive de qual seria a aceitação dos usuários para utilização da ferramenta. Para finalizar, foram feitos testes para verificar a qualidade do *feedback*.

Estas informações foram coletadas através de questionários, elaborados de acordo com os conceitos descritos por Robson [1993] e que possuíam duas partes principais, uma de dados de identificação tais como, nome, idade, tempo de experiência com computação, se já era usuário de *chats* ou não, e uma segunda parte onde deveriam responder sobre a iteração com o sistema, se acharam o sistema rápido, se foi fácil de utilizar, se o *feedback* era interessante e se o uso da ferramenta havia motivado a aula. Ainda existia espaço para sugestões de melhorias, inclusão de novas funcionalidades e reclamações. Os questionários respondidos pelos professores continham perguntas diferenciadas, sobre a ferramenta, para verificação de sua coerência perante as observações em sala de aula, utilidade de resposta, entre outras.

A ferramenta foi considerada de grande auxílio para o professor e surtiu muito interesse por parte dos alunos que relataram que a Oxentchê-chat proporciona um maior entendimento de suas ações em um *chat*, auxiliando no processo de colaboração e melhorando a aprendizagem.

5. Conclusões e Trabalhos Futuros

Através dos testes realizados nas universidades reais e dos testes feitos pelo grupo de pesquisa da UFPE percebeu-se o interesse dos alunos em utilizar a ferramenta. Um dos principais motivos admitidos foi que a Oxentchê-chat possui as funcionalidades necessárias e existentes na maioria dos *chats* disponíveis na *web*, além da análise dos diálogos. Isto que permite um monitoramento sem a necessidade do aprendizado de um sistema específico e novo, o que muitas vezes pode causar desmotivação e evasão.

A ferramenta vem sendo aprimorada de acordo com as necessidades apresentadas pelos próprios alunos e professores que a utilizam, mostrando o funcionamento do sistema em ambientes reais.

Com os resultados obtidos, é possível concluir que as Redes Neurais Artificiais e as Árvores de Decisão podem ser utilizadas para classificar diálogos educacionais e auxiliar efetivamente os professores e alunos a monitorar o processo de ensino/aprendizagem.

Como trabalhos futuros pode-se citar a continuidade de testes em ambientes de aprendizado (reais ou virtuais), além da escolha de novas abordagens para comparação

de resultados com as RNAs e ADs. Também pode-se disponibilizar a ferramenta para outros domínios tais como, comunidades virtuais de aprendizado e cursos a distância.

Referências

- Boud, D., Keogh, R. and Walker, D., (1985). “Promoting Reflection in Learning: a Model”. In Boud, D., Keogh, R. and Walker, D.(eds.) *Reflection: Turning Experience into Learning*, Kogan Page, 18-40
- Haykin, S. (1994) *Neural Networks – A Comprehensive Foundation*. Prentice-Hall. 696p.
- Marcuschi, L. A.; (2003) *Análise da Conversação* Editora Ática, 94p.
- Robson, C., *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*, Blackwell (1993)
- Robertson, J.; Good, J.; Pain, H. (1998) “BetterBlether: The Design and Evaluation of a Discussion Tool for Education”. *International Journal of Artificial Intelligence in Education*, 9, 219-236
- Russel e Norvig (2003) *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2ª Edition. 932p.
- Searle, J. (1972) “What is a speech act?”. In P. Giglioli (Ed.), *Language and Social Context*. Penguin Books Ltd., 136-154
- Siebra, S.A; Christ, C.R.; Queiroz, A.E.M.; Tedesco, P.C.A.R.; Barros, F.A. (2004) “SmartChat - An Intelligent Environment for Collaborative Discussions”. Em: 7th International Conference on Intelligent Tutoring Systems (ITS’2004)
- Soller A.; Wiebe J.; Lesgold A., (2002). “A Machine Learning Approach to Assessing Knowledge Sharing During Collaborative Learning Activities”. *Proceedings of Computer Support for Collaborative Learning 2002*, 128-137.
- Torreão, P.; Aquino, M.; Tedesco, P.; Sá, J. e Correia, A. (2004): A Methodology for the Construction of Learning Companions. Apresentado como Poster no ITS’2004.
- Vieira, A. C. H. (2004) *Classificando automaticamente diálogos colaborativos on-line com a OXEnTCHÊ-Chat* Dissertação de Mestrado em Ciência da Computação. UFPE, 128p.
- Vieira A. C. H.; Teixeira, L.; Timóteo, A; Tedesco, P. e Barros, F. (2004) “Analyzing On-Line Collaborative Dialogues: The OXEnTCHÊ-Chat”. Em: 7th International Conference on Intelligent Tutoring Systems (ITS’2004)