UMA METODOLOGIA PARA AVALIAÇÃO PERMANENTE DO ENSINO BASEADA NO SOFTWARE AVALIAFACIL E TÉCNICAS DE MINERAÇÃO DE DADOS

Débora Madonado[1], Diemison Carlos Melo[2], Bruno Gomes Haick[3], Aldebaro Klautau[4]

Abstract — *This paper presents a methodology for evaluating an undergraduate course, which was successfully implemented by the Department of Computer Engineering, at the Federal University of Para, and can be adopted by other institutions and courses. The methodology is based on two steps. The first is the periodic collection of data through the web software Avaliafacil, which is open source and available to the community. Avaliafacil allows teachers, students and staff to evaluate the course by filling in forms via the Internet. The second step is the processing of information through conventional statistics (for example, statistical correlation) and data mining techniques (decision trees, association rules, etc.), which try to discover patterns not easily identified in the collected data.*

Index Terms — evaluation methodology, data mining, learning assessing.

INTRODUÇÃO

O ensino de Engenharia da Computação apresenta diversos desafios. Um deles é a dificuldade em se manter atualizados os laboratórios de prática, em função da vertiginosa evolução da tecnologia. Não menos importante que os equipamentos e a logística é o engajamento dos docentes e outros aspectos humanísticos que impactam a qualidade do curso. Independente do aspecto em consideração, a melhoria do ensino depende de uma avaliação continuada da qualidade do curso oferecido.

Relacionadas aos fatores humanísticos, independente do curso em questão, existem diversas barreiras para que se obtenha um curso com qualidade máxima e existem diversos pontos positivos que muitas vezes não são evidentes a todos, ficando, assim, inexplorados. Avaliar o andamento do curso, as condições de ensino e a resposta dos estudantes, torna-se um fator importante para aprimorar a qualidade oferecida e formar profissionais mais capacitados e eficientes.

Nesse contexto, foi implantada uma metodologia de diagnóstico de cursos que busca identificar deficiências e qualidades, apontando as principais características do curso. Para uma melhor organização dos resultados, esta metodologia é dividida em duas partes.

A primeira etapa consiste na coleta dos dados referentes à avaliação, que é feita através do software Avaliafacil, que é de domínio público. A coleta é realizada ao término de cada período do curso. No caso da Faculdade de Engenharia da Computação da UFPA (EngComp), a coleta é realizada semestralmente. Nessa etapa, os usuários do software são divididos em três categorias: discentes, docentes e funcionários da faculdade, avaliam a infra-estrutura do curso e também as outras categorias.

A finalização do preenchimento dos formulários de avaliação dá prosseguimento à segunda etapa, pois resulta em uma base de dados, que será tratada com utilização de técnicas e algoritmos de mineração de dados, visando a eventual descoberta de padrões não identificáveis explicitamente. Essa etapa é executada pelo software que camufla o uso da API (Application Programming Interface): Weka (Waikato Environment for Knowledge Analysis)[10]. Após essa etapa, o sistema Avaliafacil disponibiliza os resultados para que sejam interpretados, visualizados e socializados com a faculdade.

O SOFTWARE AVALIAFACIL

O processo de avaliação na EngComp já vinha sendo implantado através de formulários impressos. Contudo o processo era custoso e demorado. Após levantamento de requisitos, determinou-se o desenvolvimento do software, de maneira que este sistema se tornasse simples para o usuário-avaliador, permitindo a obtenção de uma quantidade considerável de avaliadores das três categorias de maneira otimizada.

O Avaliafacil foi desenvolvido em PHP (*Hypertext Preprocessor*) [7], utiliza o MySQL [6] como sistema gerenciador de banco de dados, também utiliza SQL (*Structured Query Language*) [9] para a comunicação entre PHP e o banco de dados.

O software foi dividido em dois módulos: Administração e Avaliação. O módulo Administração corresponde ao ambiente onde poderá ser configurado o formulário. Inserções, remoções e alterações de perguntas poderão ser feitas pelo administrador do sistema, além de decidir que tipo de resposta cada pergunta terá, podendo ser de caráter qualitativo ou quantitativo. Além disso, ao final de cada avaliação, o administrador pode gerar relatórios por disciplina, além de obter conhecimentos

1 Débora Silva Madonado Cunha, Estudante de Graduação da Faculdade de Engenharia da Computação - Universidade Federal do Pará, deboracunha@ufpa.br
2 Diemison Carlos Romano de Melo, Estudante de Graduação da EngComp - Universidade Federal do Pará, dcarlos@ufpa.br
3 Bruno Gomes Haick, Estudante de Graduação da EngComp - Universidade Federal do Pará, , bghaick@ufpa.br
4 Aldebaro Barreto da Rocha Klautau Junior,Diretor da EngComp, aldebaro@ufpa.br

estatísticos sobre a avaliação já ocorrida, como quantos avaliadores participaram, quantos avaliadores deveriam participar, entre outras informações fornecidas. A Figura 1 apresenta o diagrama de casos de uso [3] do módulo Administração.

FIGURA 1
DIAGRAMA DE CASOS DE USO DO MÓDULO ADMINISTRAÇÃO.

O módulo Avaliação é destinado para os avaliadores das três categorias: discentes, docentes e funcionários da faculdade. Cada avaliador preenche formulários relacionados às disciplinas que têm vínculo e à infra-estrutura do curso, além de poder informar algum comentário sobre o aspecto que está sendo avaliado, como pode ser visto na Figura 2, que apresenta o diagrama de casos de uso do módulo Avaliação.

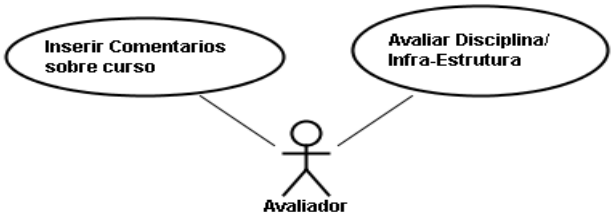


FIGURA 2
DIAGRAMA DE CASOS DE USO DO MÓDULO AVALIAÇÃO.

O software tem como importante critério a confidencialidade das respostas de cada avaliador, ou seja, a resposta de um avaliador não deve ser visualizada por nenhum usuário do sistema, não existe vínculo entre o avaliador e sua resposta. O anonimato do usuário que avaliou a disciplina é de grande importância. Para isso, modelar um banco de dados que provê a garantia desse sigilo é necessário. A opção adotada foi criar duas tabelas onde, em uma delas, "av_avaliadorXdisciplina", fornece apenas o registro do identificador do avaliador e o identificador da disciplina avaliada, para fins de controle, evitando que a disciplina seja avaliada pelo usuário mais de uma vez. Na outra tabela, "av_resposta", os campos "sim", "parcial" e "não" são atualizados de acordo com a resposta do avaliador para uma questão referente a uma disciplina. Esta organização é apresentada nas Tabelas I e II, as quais apresentam, como exemplo, o esquema de três avaliadores que respondem a três questões referentes a três disciplinas. No modelo entidade-relacionamento [4, 5] apresentado na Figura 3, mostra-se o projeto de banco de dados utilizado no software Avaliafacil.

TABELA I
TABELA AV_AVALIADORXDISCIPLINA

idAvaliador	idDisciplina
09080001101	001
09080001101	002
09080001201	001
09080001201	002
09080001201	003
09080001301	001
09080001301	003

TABELA II
TABELA AV_RESPOSTA

idDisciplina	questão	sim	parcial	não
001	1	0	2	1
001	2	1	1	1
001	3	2	1	0
002	1	2	0	0
002	2	2	0	0
002	3	1	1	0
003	1	0	1	1
003	2	2	0	0
003	3	0	0	2

Outro requisito é fazer com que um maior número de avaliadores participe do processo de avaliação. Para isso o Avaliafacil é elaborado de maneira que permita uma fácil utilização, onde poucos passos fossem necessários para finalizar a avaliação. A Figura 4 indica os passos dados por um avaliador, mostrando o processo de avaliação via Avaliafacil. Estes passos totalizam três, autenticação do usuário (Passo 1), escolha do aspecto que será avaliado (Passo 2), e a avaliação através do formulário (Passo 3).

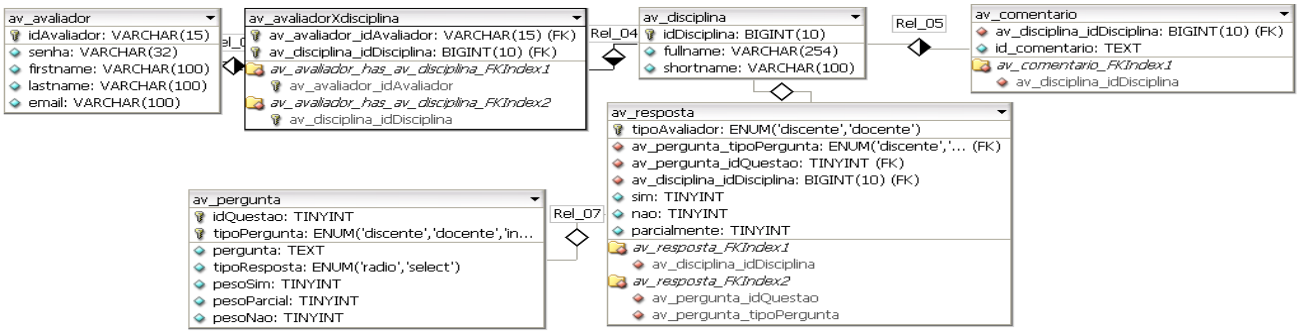


FIGURA 3
MODELO ENTIDADE RELACIONAMENTO DO SOFTWARE AVALIAFACIL.

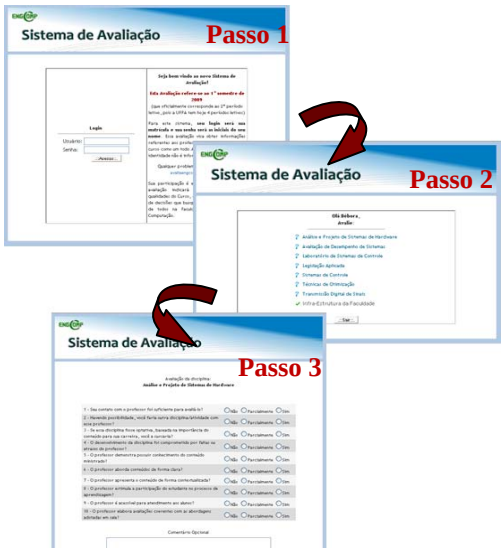


FIGURA 4
ESQUEMA DO PROCEDIMENTO REALIZADO NO SOFTWARE **AVALIAFACIL** NO PROCESSO DE AVALIAÇÃO POR PARTE DOS AVALIADORES

O PROCESSO DE AVALIAÇÃO

Para a avaliação da EngComp, foram selecionadas 10 perguntas para os discentes avaliarem os docentes, 9 perguntas para os docentes avaliarem os discentes e 10 perguntas referentes a infra-estrutura da faculdade, onde discentes, docentes e funcionários, avaliam a coordenação da faculdade.Os discentes avaliam cada docente que tem vinculo, já os docentes avaliam a turma formada pelos discentes.

A avaliação ocorreu em um período de 20 dias, e obteve-se participação de aproximadamente 70% da comunidade, o que fornece uma aceitável base de dados para interpretações estatísticas através de técnicas convencionais de mineração de dados.

TÉCNICAS DE MINERAÇÃO DE DADOS

Para uma exploração adequada do conjunto de dados obtido através do processo de avaliação pelo software Avaliafacil, utilizamos a mineração de dados [8] afim de extrair e evidenciar padrões nesta base de dados, fornecendo informações interessantes para o aperfeiçoamento da qualidade de ensino da EngComp.

Inicialmente, realiza-se uma rotina de pré-processamento dos dados, na tentativa de corrigir inconsistências.

Com este intuito, utilizou-se o pacote Weka [10], que é formado por um conjunto de implementações de algoritmos de diversas técnicas de mineração de dados [1]. Este software é licenciado pela *General Public License*, e possui uma série de heurísticas relacionadas a classificação, regressão, clusterização, regras de associação e visualização, e também fornece a visualização das arvores de decisões. Além disso, este software possui uma interface de fácil utilização e seus algoritmos disponibilizam relatórios com dados analíticos e estatísticos do domínio minerado.

Um arquivo no formato *.arff* é necessário para a leitura pelo Weka. Sendo assim, o arquivo é organizado basicamente em duas partes:

- Lista de atributos: onde se define o tipo de atributos e o valor que ele pode representar. O número total de atributos é 41. Cada pergunta pode ser respondida como "sim", "não", "parcial". Por exemplo, para a primeira pergunta, os parâmetros P1Sim, P1Parcial, P1Nao indicam a porcentagem em cada categoria. Como a avaliação do docente tem 10 perguntas (vide Anexo), a quantidade de atributos referentes a essas perguntas soma 30. Além desses 30 atributos, incluiu-se 10 outros: o nível de formação do docente (doutorado, mestrado, especialização, graduação), tipo de atuação (permanente ou substituto), tempo de atuação do docente, se o docente atua em algum programa de pós-graduação, quantidade de discentes que avaliou o curso, quantidade de discentes matriculados, quantidade de discentes aprovados no curso, reprovados por falta, reprovados por desempenho insuficiente e reprovados sem avaliação na disciplina.. O ultimo parâmetro, que completa 41 atributos no .arff, é um número chamado "resultado", o qual é obtido a partir dos dados da avaliação e que objetiva fazer um "ranking" por disciplina. O "resultado" M_p é usado para atribuir um conceito ao docente como ótimo, bom, regular ou insuficiente, como mostrado na Tabela III.

- Dados: para cada disciplina, correspondendo a uma linha no arquivo .arff, informa-se os 41 valores referentes a cada atributo (coluna). São esses os registros que serão minerados.

TABELA III
TABELA DE DEFINIÇÃO DO LIMIAR PARA CONCEITO DO DOCENTE

Resultado da Avaliação	Conceito
$M_p >= 43$	Ótimo
$43 > M_p >= 37$	Bom
$37 > M_p >= 30$	Regular
$M_p < 30$	Insuficiente

Para os dados em questão o valor máximo do "resultado" M_p foi igual a 50, e o mínimo igual a 16. Descreve-se a seguir como se obter M_p. Para cada uma das 10 perguntas é atribuído um peso ω_n onde *n* refere-se à *n*-ésima pergunta, de maneira que se possa indicar quais as mais importantes. A porcentagem das respostas nas categorias (chamadas aqui tipos) "sim", "não" e "parcialmente" para a *n*-ésima pergunta é expressa por R_n. Para cada disciplina e cada um dos três tipos (N_{sim} , $N_{parcial}$, $N_{não}$) é calculado

$$N_{tipo} = (\omega_{1,tipo}.R_{1,tipo} + \omega_2.R_{2,tipo} + \omega_3.R_{3,tipo} + \ldots + \omega_n.R_{n,tipo})/\Sigma\, \omega_n \quad (1)$$

onde *n*=10 (número de perguntas). Após a obtenção de 3 valores de *N* para cada disciplina, calcula-se o resultado que será usado para o ranking:

$$M_p = (\mu_{sim}.\ N_{sim} + \mu_{parcial}.\ N_{parcial} + \mu_{nao}.N_{não})/\Sigma\ \mu \quad (2)$$

onde os pesos μ são valores que permitem converter as porcentagens de sim, não, parcialmente em um único número M_p. Para o presente trabalho, usou-se ω_n=1 (mesmo peso para todas as 10 perguntas) e μ_{sim}=3, $\mu_{parcial}$=2, μ_{nao}=1 (nota-se que o importante é a diferença entre os diversos μ e, por exemplo, μ_{sim}=1, $\mu_{parcial}$=0, μ_{nao}=-1 conduziriam ao mesmo resultado).

Após a conversão para números, os mesmos podem ser tratados por técnicas convencionais. De forma similar, após gerar o .arff, utilizam-se os algoritmos do Weka para mineração de dados, como exemplificado na próxima seção.

RESULTADOS

Como exemplo de tratamento convencional, a Figura 5 apresenta a média geral das disciplinas em relação às perguntas, onde foram analisados os valores para a classificação do docente.

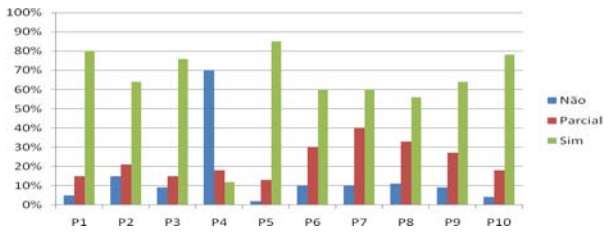


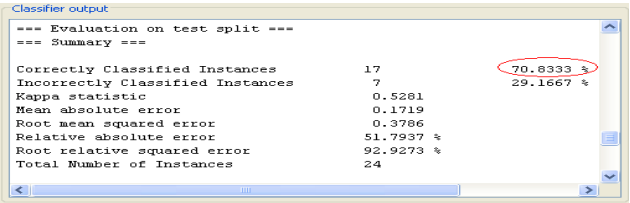
FIGURA 5
A ABSCISSA INDICA AS 10 PERGUNTAS E A ORDENADA INDICA O RESULTADO GERAL PARA OS 3 TIPOS "SIM", "NÃO", "PARCIALMENTE".

Algoritmo Utilizado para a Obtenção de Padrões

Como exemplo de mineração de dados, serão mostrados os resultados utilizando-se o algoritmo de classificação J48, que gera uma árvore de decisão induzida a partir de dados de treinamento [2, 10].

Para isso, os parâmetros de projeto da árvore foram ajustados com o intuito de aprimorar os resultados obtidos (etapa chamada de "model selection").

A base de dados foi dividida em duas, sendo a primeira parte utilizada para treinamento e a segunda para teste. Este ajuste fornece uma taxa de acerto de 70%, que pode ser considerada uma acuracidade suficiente para o estudo de padrões de interesse. A Figura 6 apresenta a resposta do algoritmo à base dados e critérios definidos.



```
Classifier output
=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances          17               70.8333 %
Incorrectly Classified Instances         7               29.1667 %
Kappa statistic                          0.5281
Mean absolute error                      0.1719
Root mean squared error                  0.3786
Relative absolute error                 51.7937 %
Root relative squared error             92.9273 %
Total Number of Instances               24
```

FIGURA 6
EXIBIÇÃO DO WEKA EXPLORER ONDE SÃO APRESENTADAS AS RESPOSTAS DO ALGORITMO

A Figura 7 mostra a árvore de decisão gerada a partir dos critérios estabelecidos acima.

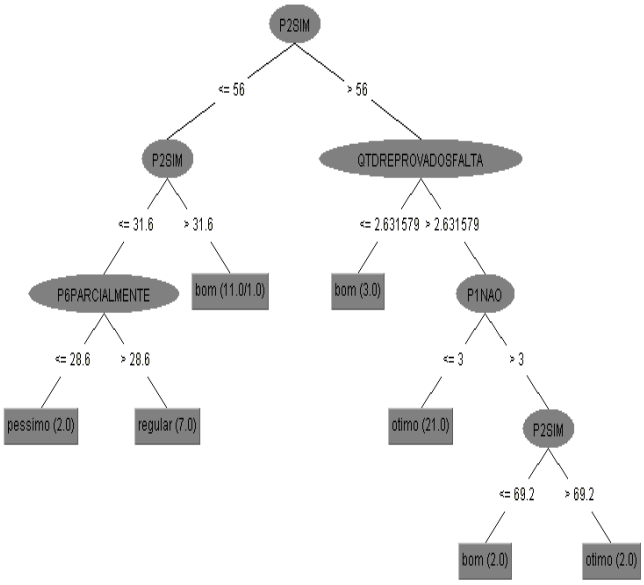


FIGURA 7
ÁRVORE DE DECISÃO GERADA PELO ALGORITMO J48

A análise da árvore de decisão mostrada Figura 7 permite avaliar os principais critérios que conceituam um docente como ótimo, bom, regular ou insuficiente.

O item P2 corresponde à pergunta onde é avaliado se os discentes gostariam de ter outra disciplina com o docente em questão. Nesta árvore, a resposta afirmativa, representada pelo nó P2Sim, é a raiz da árvore, que define o inicio de todos os caminhos, ou seja, inicia todas as regras de classificação.

Para uma dada disciplina, caso P2Sim esteja entre 31% e 56% de respostas afirmativas, o docente é conceituado como **bom**, senão (caso P2Sim obtenha menos de 31.6% de respostas afirmativas) o conceito do docente fica dependendo da porcentagem do item P6, que corresponde à pergunta onde é avaliada a clareza nas explicações do professor em sala de aula. Nesta decisão, a resposta que define o conceito atribuído ao docente é a porcentagem de respostas "parcialmente", podendo conduzir à sugestão do conceito **insuficiente** ou **regular**.

Entretanto se P2Sim tem uma porcentagem maior que 56%, verifica-se a porcentagem de alunos reprovados por índice de freqüência insuficiente. Se este índice for menor que 2.63%, o docente é conceituado como **bom**, porém se o índice for maior, a decisão é realizada em cima da porcentagem de discentes que se dizem aptos a avaliar o docente, se a porcentagem de discentes não aptos for menor ou igual a 3%, o docente é avaliado como **ótimo**, caso contrário, um último teste é realizado referente a P2Sim novamente, onde se a porcentagem deste for superior a 69.2%, o docente é avaliado como **ótimo**, senão, é avaliado como **bom**. Como fato interessante, nota-se que a árvore identificou a relação que há entre o

número de reprovados e o conceito final dado aos discentes à disciplina.

CONCLUSÕES

Neste artigo, foi apresentado o software Avaliafacil que tem como principal propósito adquirir dados de um ciclo de avaliação através de formulários online, a fim de gerar uma base de dados consistente, sendo um sistema de uso amigável. Uma alternativa para conceituar os docentes de acordo com os resultados da avaliação foi proposta. Além disso, técnicas de mineração de dados foram apresentadas, através do Weka, com o intuito de fornecer padrões coerentes para conceituar o docente, através do algoritmo de classificação comumente utilizado, o J48.

Os padrões revelados apontam os principais critérios na avaliação de um docente como sendo:

- Aceitação dos discentes em ter outra disciplina ministrada pelo docente em questão (P2);
- Quantidade de alunos reprovados por freqüência insuficiente;
- Clareza da abordagem da disciplina pelo docente (P6);
- Quantidade de discentes que se revelam aptos a avaliar o docente (P1).

Estes critérios auxiliam na compreensão dos conceitos dados pela avaliação, objetivando o aprimoramento da qualidade de ensino da EngComp.

O mesmo procedimento descrito para a avaliação e classificação do docente pelos discentes, é utilizado para a avaliação e classificação dos discentes pelo docente, e infra-estrutura por ambos docentes e discentes.

AGRADECIMENTOS

Agradecemos ao Prof. Claudomir Cardoso de Carvalho Junior pelo auxílio no desenvolvimento deste trabalho.

REFERÊNCIAS

[1] Amo, S., "Técnicas de Mineração de Dados" – Sociedade Brasileira de Computação, Universidade Federal da Bahia. (Org.). Jornadas de Atualização em Informática. Salvador: Universidade Federal da Bahia, 2004, v. 2, p. 195-236.

[2] Almeida, L. M., Padilha, T.P.P., Oliveira, F. L., Previero, C. A.,"Uma Ferramenta para Extração de Padrões" – REIC. Revista Eletrônica de Iniciação Científica, v. III, p. 7, 2003.

[3] Booch, G., Rumbaugh, J. , Jacobson, I., "UML : Guia do Usuário", Campus, 2ª Edição, 2005 .

[4] DBdesigner – http://fabforce.net/dbdesigner4. Acesso em: 26 de outubro de 2009.

[5] Heuser, C. A., "Projeto de Banco de Dados", Bookman, 6ª Edição, 2009.

[6] MySQL Open Source Database. http://www.mysql.com. Acesso em:26 de outubro de 2009.

[7] PHP – Hypertext Preprocessor. http://www.php.net. Acesso em: 26 de outubro de 2009.

[8] Silva, M. P. S., "Mineração de Dados: Conceitos, Aplicações e Experimentos com Weka", Livro da Escola Regional de Informática Rio de Janeiro, Espírito Santo, 2004.

[9] SQL Database Reference Material. http://www.sql.org. Acesso em: 26 de outubro de 2009.

[10] University of Waikato.Weka – Machine Learning Software in Java. http://www.cs.waikato.ac.nz/ml/weka/uml. Acesso em: 26 de outubro de 2009.

ANEXO

As 10 perguntas utilizadas foram:

P1 – Seu contato com o professor foi suficiente para avaliá-lo?

P2 – Havendo possibilidade, você faria outra disciplina/atividade com esse professor?

P3 – Se essa disciplina fosse optativa, baseada na importância do conteúdo para sua carreira, você a cursaria?

P4 – O desenvolvimento da disciplina foi comprometido por faltas ou atrasos do professor?

P5 – O professor demonstra possuir conhecimento do conteúdo ministrado?

P6 – O professor aborda conteúdos de forma clara?

P7 – O professor apresenta o conteúdo de forma contextualizada?

P8 – O professor estimula a participação do estudante no processo de aprendizagem?

P9 – O professor é acessível para atendimento aos alunos?

P10 – O professor elabora avaliações coerentes com as abordagens adotadas em sala?