



S.E.A.S.

Student Evasion Analysis System

Arthur Vianna Landeo
Cassiano Ricardo S C Filho
Tatiane Portela Medeiros

Orientador: Prof. Dr. Alexander Robert Kutzke.



Roteiro

- Introdução
- Objetivo
- Data mining, scripts classificadores e predição
- Desenvolvimento
- Apresentação do sistema
- Conclusão



Introdução

- Número de pessoas que concluem o ensino superior é baixo se comparado ao número de pessoas matriculadas;
- Evasão escolar: “Medida total do número de alunos que tendo entrado em um determinado curso de ensino superior, não obteve diploma no final do curso.” (FILHO, 2007)
- Retenção escolar: “Situação em que um estudante do ensino superior demanda um tempo superior ao previsto do calendário pedagógico para concluir o curso em que se matriculou.” (MEC/SEF, 1997)
- No TADS, de 2005 a 2018, de 568 alunos que concluíram o curso, apenas 157 conseguiram se formar no tempo previsto, tem-se um índice de retenção que chega a 72%;
- Dificuldade dos docentes em acompanhar os alunos continuamente durante o curso.



Objetivos

- Modelar e desenvolver um sistema que auxilie o corpo docente a visualizar e interpretar dados referentes à evasão e retenção dos estudantes do TADS;
- Desenvolver um algoritmo que utiliza inteligência artificial para compor uma análise dos dados do corpo discente;
- Disponibilizar esta análise e a predição gerada para o corpo docente utilizando uma interface web acessível e de fácil entendimento;



Fundamentação

Knowledge Discovery in Databases (KDD) :

- Identificação de padrões
- Dividido em etapas
 - Limpeza de dados
 - Inserção de dados
 - Seleção de dados
 - Transformação de dados
 - Data mining
 - Apresentação



Fundamentação

- Data warehouse
 - Estrutura multidimensional
 - Tabelas fato
 - Tabelas dimensão



Fundamentação

Machine learning - Aprendizado supervisionado

- K-nearest neighbor
- Decision Tree
- Naive bayes
- Multi-layer perceptron (Redes neurais)



Organização

- Reuniões digitais semanais
- Focos específicos designados aos membros
- Documentação/Versionamento em nuvem
- Git flow de feature branch



Tecnologias utilizadas

- Python 3
- Scikit-learn
- Django
- Mysql 5.7
- Jinja2
- WebApp Admin LTE 2: Bootstrap 3 e JQuery
- Highcharts.js
- Docker

Dados disponibilizados

- Foi disponibilizado um arquivo csv contendo dados anonimizados de todos que cursaram o TADS desde 2005, até o primeiro semestre de 2018.

COD_CURSO	ID_ALUNO	MATR_ALUNO	FORMA_INGRE_TAB	FORMA_INGRE_ITEM	ANO_INGRESSO	PERIODO_INGRE_TAB	PERIODO_INGRE_ITEM	FORMA_EVASAO_TAB	FORMA_EVASAO_ITEM	ANO_EVASAO	PERIODO_EVA_TAB	PERIODO_EVA_ITEM	CONC
42A	156073	02a11a6538fb00434	612	38	2007	608	201	613		4	2011	608	100
42A	156073	02a11a6538fb00434	612	38	2007	608	201	613		4	2011	608	100
42A	156073	02a11a6538fb00434	612	38	2007	608	201	613		4	2011	608	100
42A	156073	02a11a6538fb00434	612	38	2007	608	201	613		4	2011	608	100
42A	156073	02a11a6538fb00434	612	38	2007	608	201	613		4	2011	608	100
42A	156073	02a11a6538fb00434	612	38	2007	608	201	613		4	2011	608	100



Inserção dos dados

- Dois bancos de dados construídos;
- Primeiro banco guarda os dados dos alunos que ainda estão matriculados no curso;
- Nessa importação ocorre a etapa do KDD que consiste na limpeza dos dados;
- Dados dos alunos que já evadiram estão no segundo banco de dados;
- Construído seguindo o conceito de Data Warehouse;
- Nessa fase iniciou-se a etapa do KDD que consiste na integração dos dados;
- No segundo banco é realizado os cálculos para a criação dos novos dados sobre o aluno.



Testes de predição

- Várias baterias de testes :
 - Seleção de atributos
 - Normalização
 - Data mining
- Resultado final :
 - MLP
 - Percentagem de reprovações, IRA

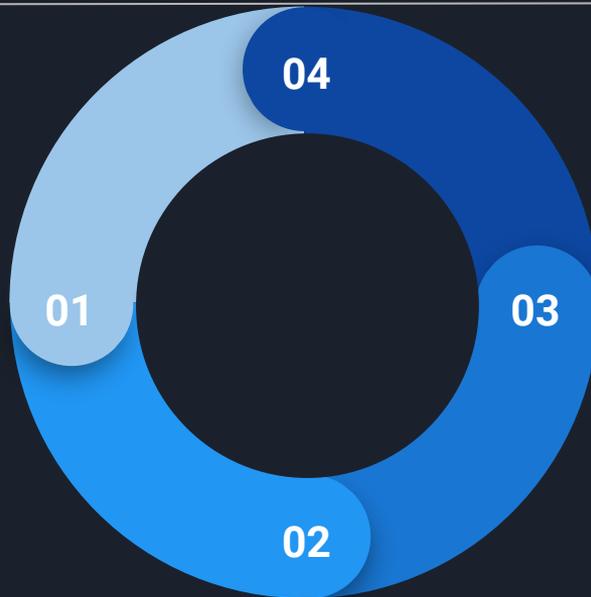
Fluxo de importação de dados

01 Inserção dos dados

Atualização dos bancos de dados

02 Treinamento do modelo

Atualizar o treinamento do modelo



04 Apresentação

Apresentação dos dados ao usuário

03 Realizar novas previsões

Preparar dados de alunos ainda matriculados e utilizá-los no modelo para criação de novas previsões

Apresentação do sistema



Conclusão

- O SEAS atingiu o objetivo proposto.
- Limitações:
 - No mínimo um semestre precisa ser feito;
 - Inserção de dados precisa ser feita no mesmo padrão do modelo disponibilizado.
- Futuro:
 - Novas predições;
 - Sistema de queries para visualização de dados.



Obrigado!





Referências

FILHO, Roberto Leal Lobo e Silva. **A evasão no ensino superior brasileiro**. Cadernos de Pesquisa, [s. l.], 1 set. 2007. Disponível em:
http://publicacoes.fcc.org.br/ojs/index.php/cp/article/view/346/350secaio.com/personal/TC/enegep2008_TN_STO_078_545_11614.pdf. Acesso em: 22 nov. 2019.

MEC/SEF. Secretaria de Educação Fundamental. **Parâmetros curriculares nacionais: introdução aos parâmetros curriculares nacionais**. Secretaria de Educação Fundamental, Brasília, 1997. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/livro01.pdf>. Acesso em: 25 nov. 2019.