

Experimento

- Universidade Federal de Pelotas – 2013
- Cursos a distância
- Turmas duram 7 semanas

Base de dados

Tabela 1. Total de alunos em cada curso e semestre

Curo	Semestre	Alunos	Aprovados	Reprovados
Educação do Campo (CLEC)	1º	133	103	30
	2º	94	88	6
Pedagogia (CLPD)	1º	196	137	59
	2º	172	119	53

- Turmas duram 7 semanas

Informação utilizada

- Contagem de interações
 - Desconsidera que atividade é realizada
 - Total de 835 mil interações
- Rótulos
 - Aprovado
 - Reprovado
- Objetivo
 - Predizer o rótulo com a maior antecedência

Vetor de características

- Vetor com número de interações em cada semana

80						
----	--	--	--	--	--	--

80	90					
----	----	--	--	--	--	--

80	90	77				
----	----	----	--	--	--	--



80	90	77	99	112	69	50
----	----	----	----	-----	----	----

Treinamento

- Entre turmas
 - Divide uma turma em 2 (A e B)
 - A é usado como treinamento e B para teste
- Entre semestre
 - A turma do primeiro semestre é usada para treinamento
 - A turma do segundo semestre é usada para teste

Atributos derivados

Tabela 2. Conjunto de Atributos

Atributo	Descrição
Situação Acadêmica	Situação final do aluno na disciplina (rótulo do exemplo)
Interações por Semana (1 até 7 semanas)	Número de interações na semana.
Média	Média do total de interações pelo número de semanas.
Mediana	Mediana do conjunto de interações por semana.
Semanas Zeradas	Número de semanas com zero interações.
Média da Diferença	Média da diferença entre a semana i e a semana $i+1$.
Razão com Professores	Razão entre o total de interações do aluno e dos professores.
Razão com Tutores	Razão entre o total de interações do aluno e dos tutores.
Fator de Empenho (1 até 7 semanas)	Razão entre as interações da semana do aluno e a média de interações da turma naquela semana.

Software

- WEKA

Resultados

Tabela 3. Precisões para classificação de reprovados – experimento sem utilização de atributos

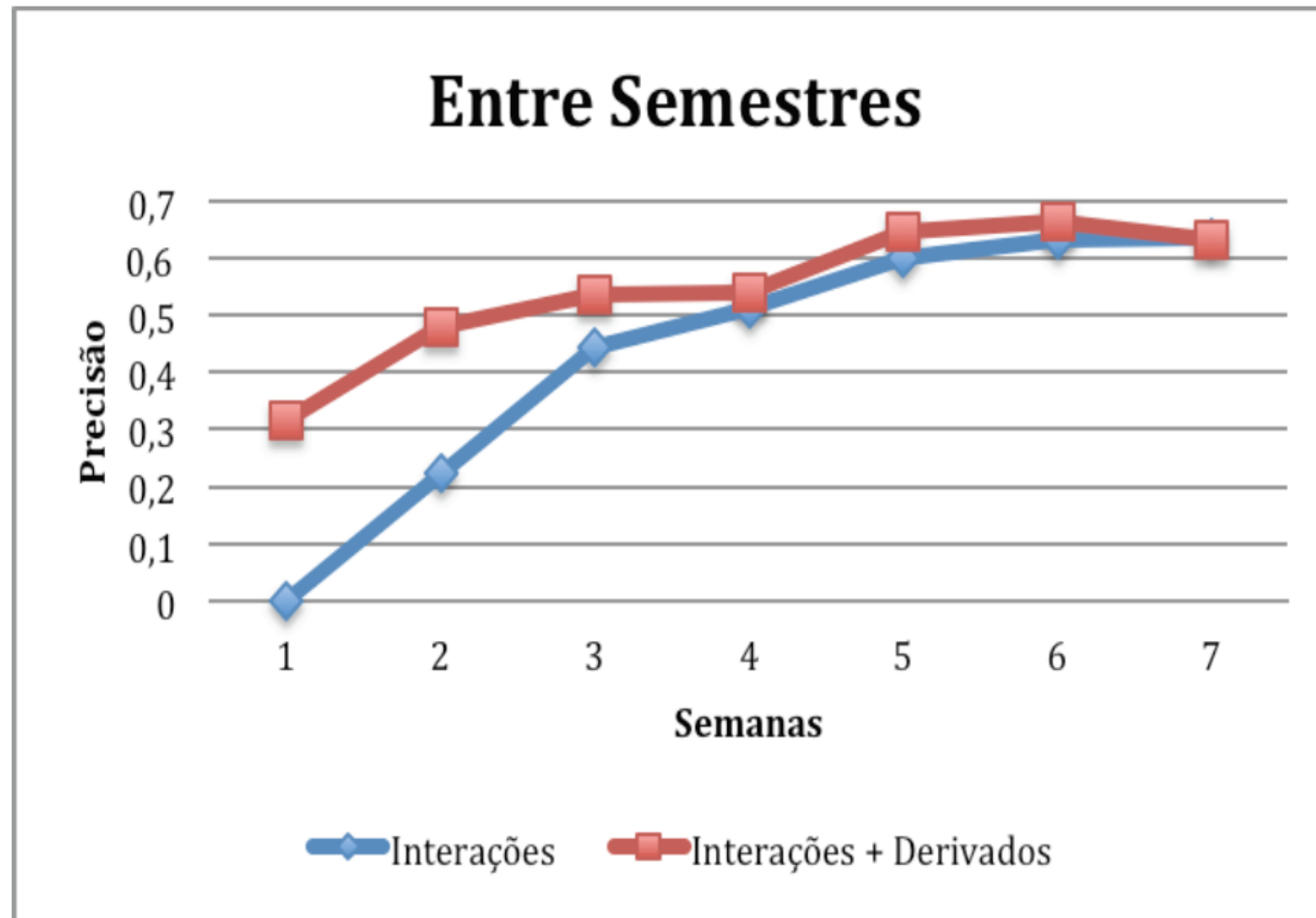
Caso	Modelo	Semana						
		S1	S2	S3	S4	S5	S6	S7
Entre Semestres	Rede Bayesiana	0,0	0,22	0,44	0,51	0,60	0,63	0,64
	Rede Neural	0,0	0,16	0,33	0,45	0,49	0,51	0,51
	J48	0,0	0,21	0,37	0,40	0,52	0,56	0,56
	Floresta Aleatória	0,1	0,33	0,36	0,40	0,48	0,51	0,54
Entre Turmas	Rede Bayesiana	0,0	0,24	0,45	0,51	0,58	0,62	0,66
	Rede Neural	0,0	0,14	0,33	0,46	0,49	0,53	0,56
	J48	0,0	0,20	0,39	0,42	0,51	0,59	0,60
	Floresta Aleatória	0,1	0,28	0,33	0,41	0,50	0,56	0,61

Resultados

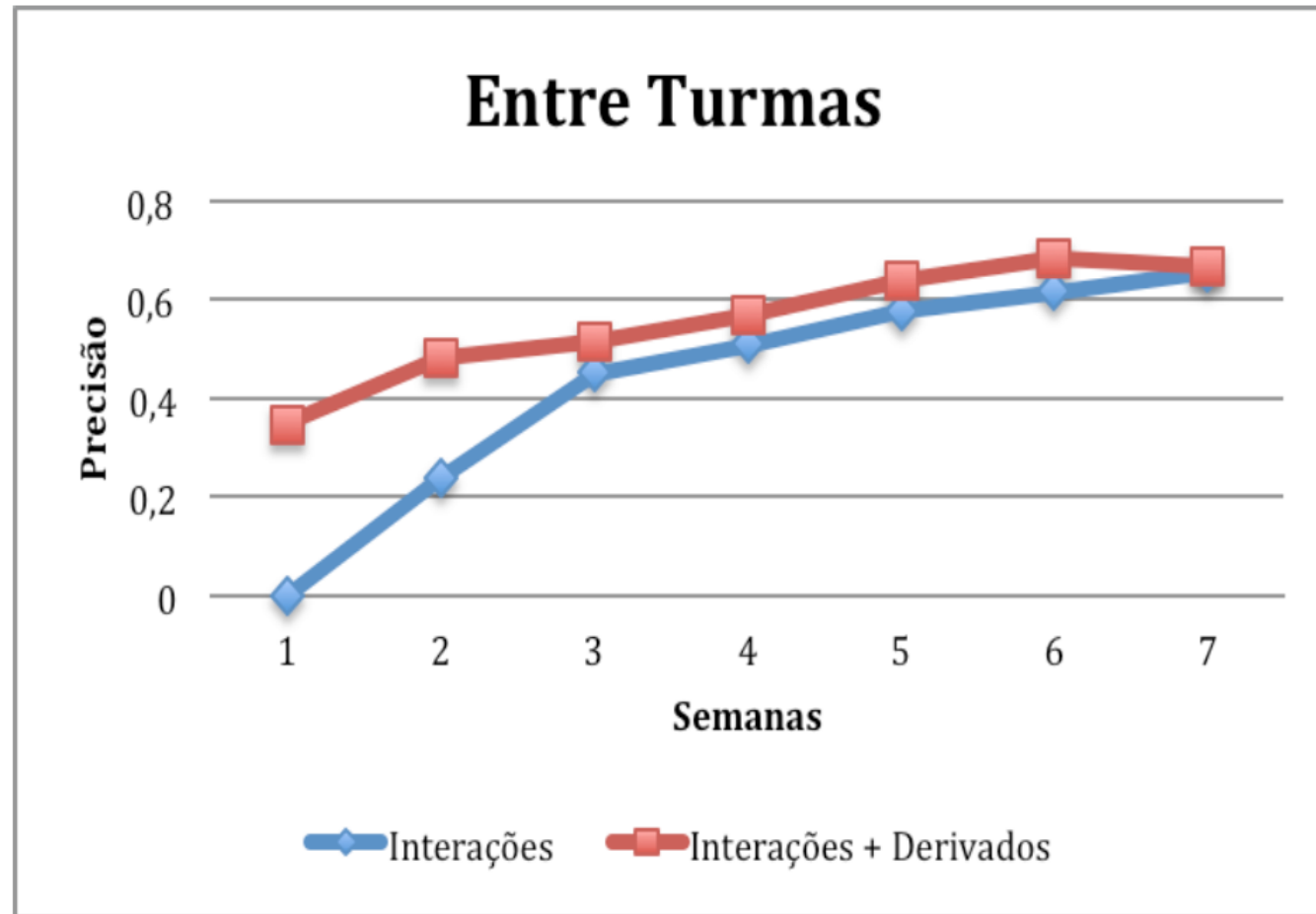
Tabela 4. Precisões para classificação de reprovados – experimento com a utilização de atributos

Testes	Modelo	Semana						
		S1	S2	S3	S4	S5	S6	S7
Entre Semestres	Rede Bayesiana	0,32	0,48	0,54	0,54	0,65	0,66	0,63
	Rede Neural	0,07	0,22	0,38	0,42	0,46	0,55	0,58
	J48	0,02	0,25	0,34	0,37	0,48	0,50	0,53
	Floresta Aleatória	0,06	0,26	0,30	0,37	0,41	0,47	0,46
Entre Turmas	Rede Bayesiana	0,35	0,48	0,52	0,57	0,64	0,68	0,67
	Rede Neural	0,08	0,24	0,41	0,44	0,49	0,57	0,59
	J48	0,05	0,27	0,36	0,39	0,49	0,57	0,64
	Floresta Aleatória	0,11	0,27	0,33	0,41	0,45	0,54	0,51

Resultados



Resultados



(b) Comparação entre turmas

Resultados com dados derivados

Figura 2. Comparação de precisão por semana utilizando Rede Bayesiana e todos atributos disponíveis

