

Analysis of LMS Data for Web-Based College Courses

Larry Dugan

Finger Lakes Community
College

Chris Mitchell

North Carolina State
University

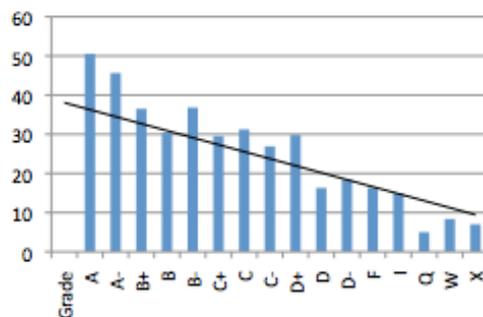
Dataset

- Predictors
 - Date Registered for Course
 - Online Credits Taken
 - Age
 - Gender
 - Number of Student Posts
 - Number of Instructor Posts
- To be predicted
 - Final Grade in Course

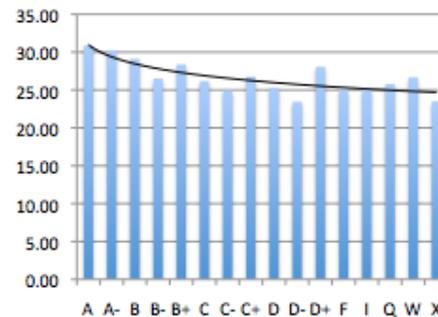
Breakout By Grade

	Values				
Row Labels	Average of NumPosts2	Count of FinalGrade	Average of Age	Average of TOTAL.CREDS	Average of RegDaysEarly
A	50.55	396.00	30.87	10.20	60.95
A-	45.62	165.00	30.24	10.17	60.88
B	30.39	191.00	29.12	10.76	65.07
B-	36.82	92.00	26.52	10.85	58.89
B+	36.49	158.00	28.37	10.50	71.88
C	31.24	114.00	26.16	10.88	58.58
C-	26.92	49.00	24.90	10.94	65.82
C+	29.55	78.00	26.76	11.60	68.50
D	16.30	56.00	25.32	10.77	55.89
D-	18.45	20.00	23.45	11.35	58.95
D+	29.79	29.00	28.07	11.10	51.62
F	16.34	200.00	25.02	9.84	45.20
I	14.93	15.00	25.07	9.60	25.40
Q	4.99	332.00	25.79	4.30	37.90
W	8.38	175.00	26.67	4.59	49.45
X	7.00	4.00	23.50	0.00	13.25
(blank)					
Grand Total	28.19	2074.00	27.67	8.97	55.52

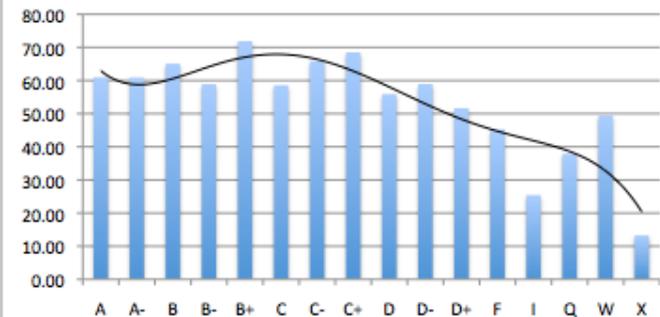
Grades by Posts



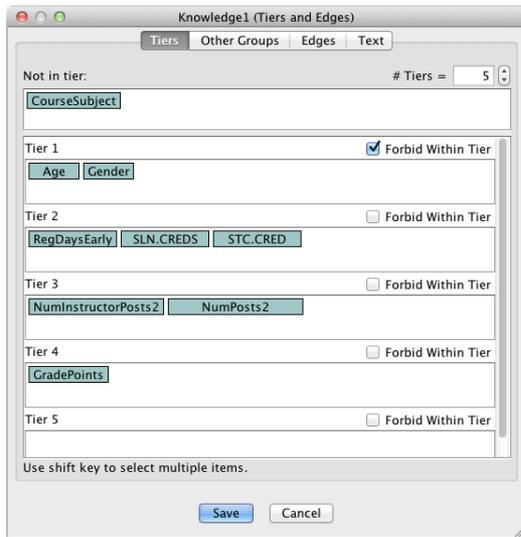
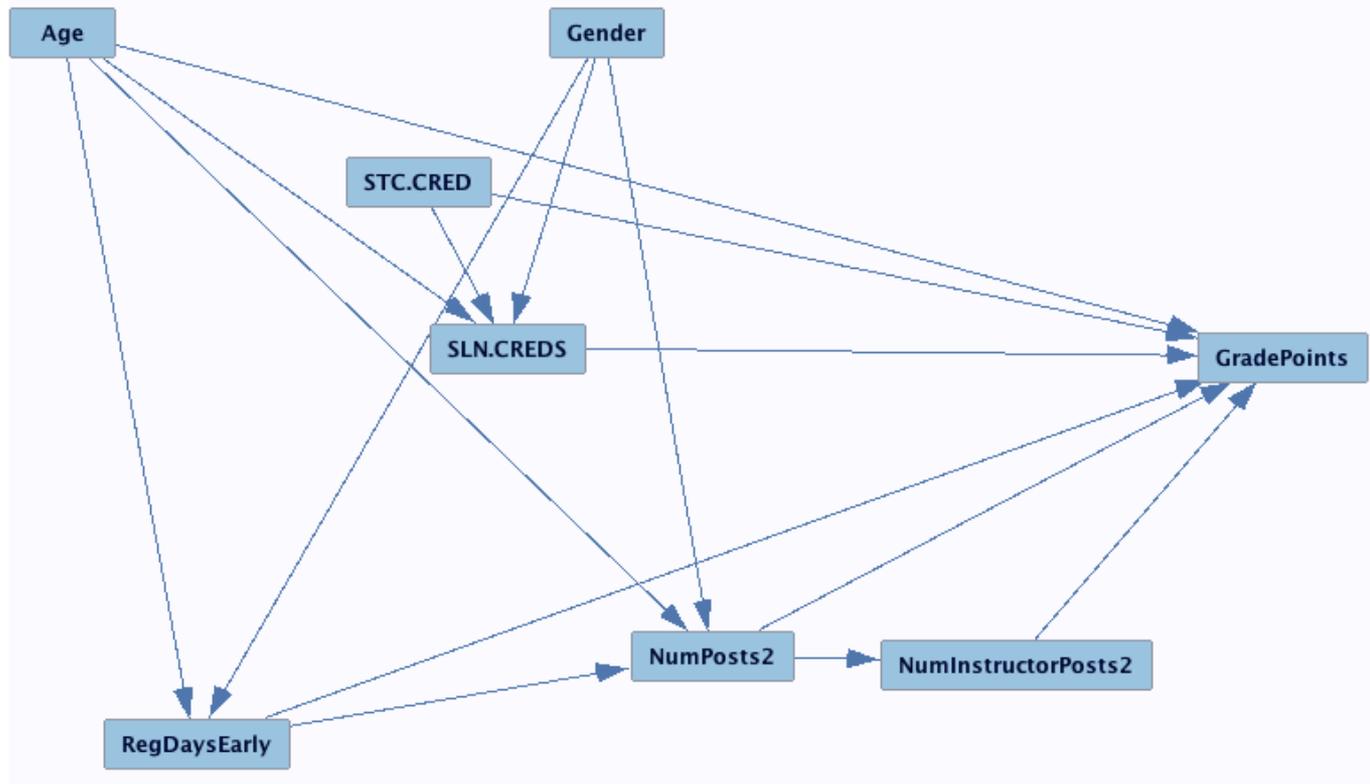
Average Age



Registration Date

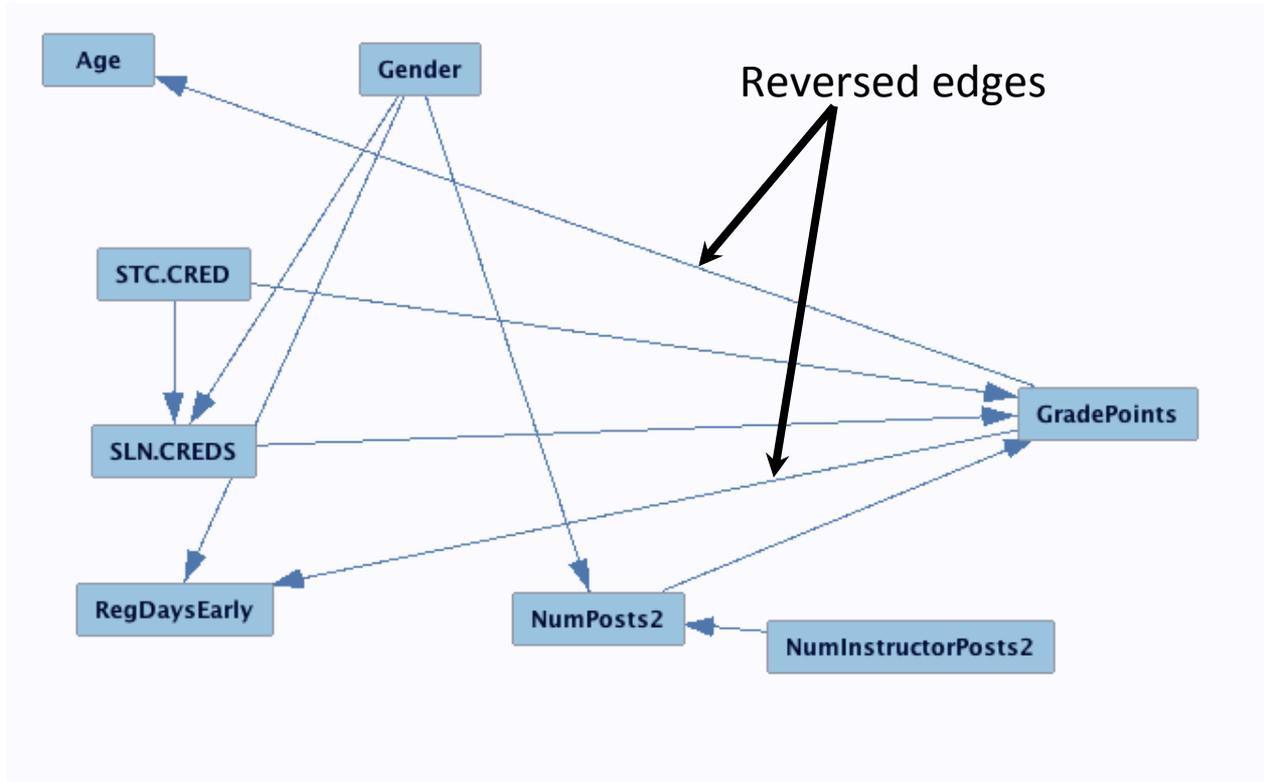


Causal Model



With Prior Knowledge

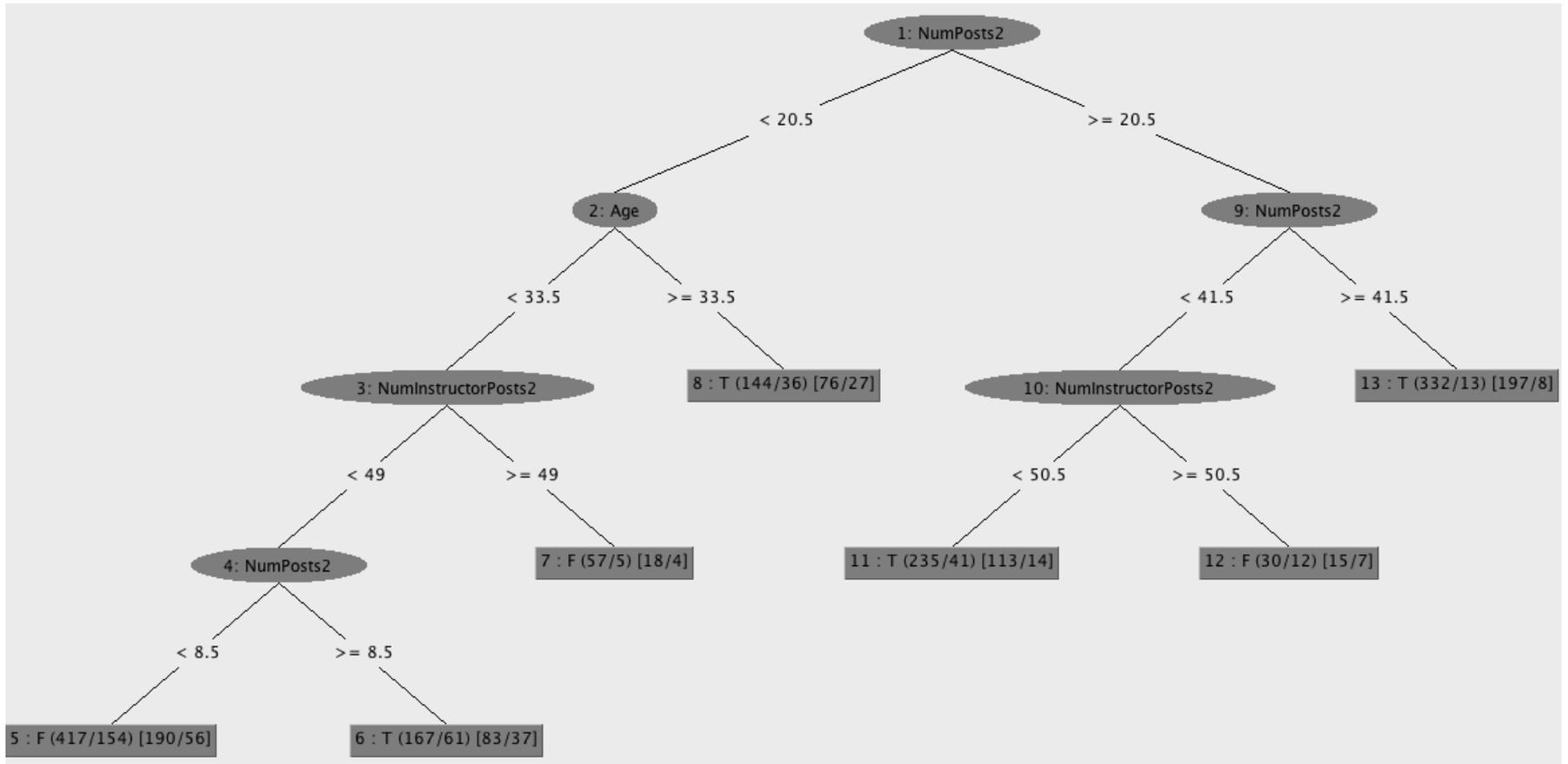
Causal model



Without Prior Knowledge

Predicting Pass/Fail

- Baseline (majority class): 65% accuracy
 - 1348 passing of 2074
- Accuracy obtained from various decision tree algorithms was around 75%
 - Number of posts to class forum appears to be best predictor of course success



REPTree Classifier Output

Accuracy: .759

Precision: .756

Recall: .759

Best-First Decision Tree Output:

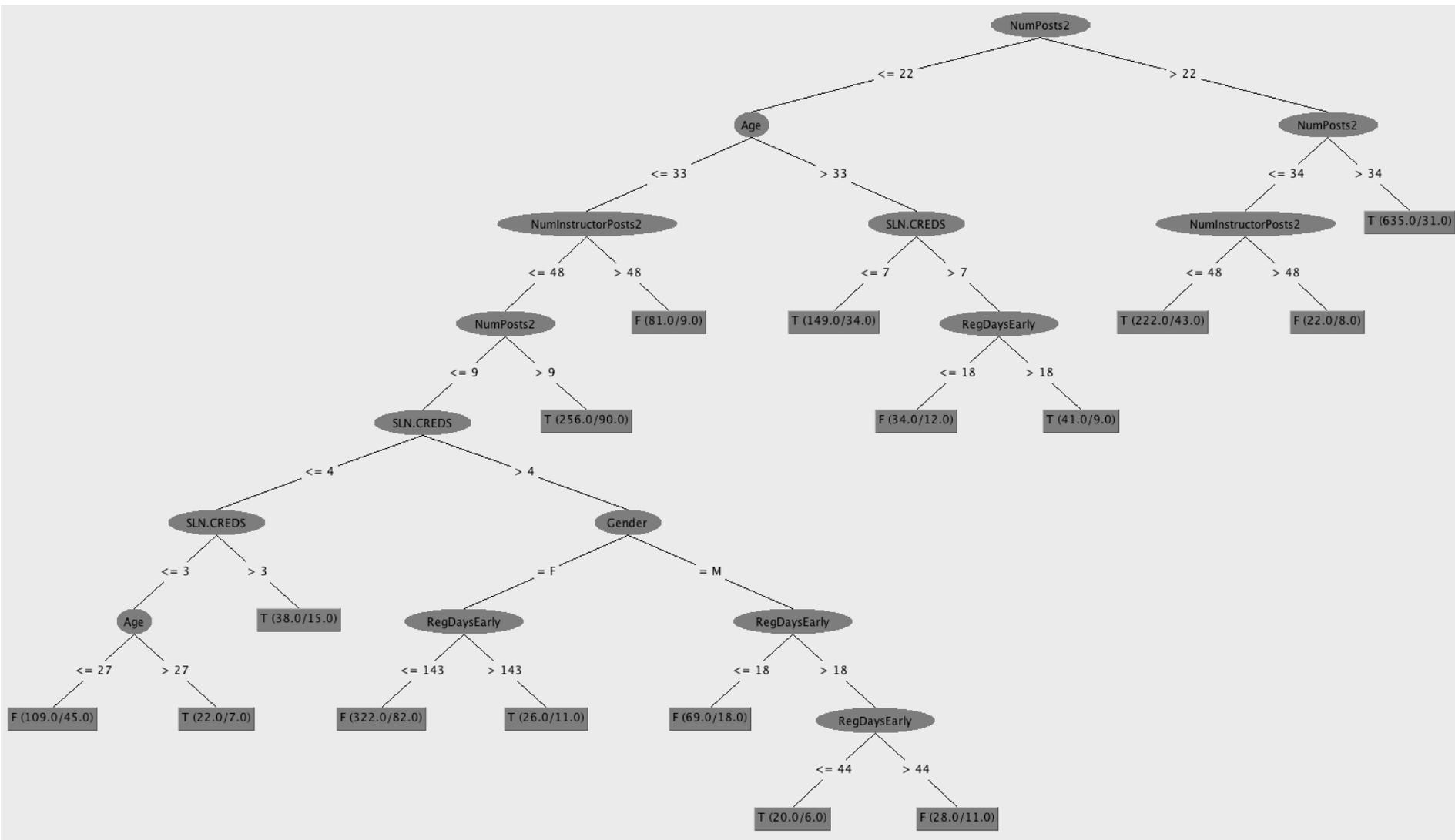
Accuracy: .761

Precision: .756

Recall: .761

```

NumPosts2 < 21.5
|   Age < 33.5
|   |   NumPosts2 < 8.5: F(433.0/215.0)
|   |   NumPosts2 >= 8.5
|   |   |   NumInstructorPosts2 < 49.0: T(165.0/101.0)
|   |   |   NumInstructorPosts2 >= 49.0: F(34.0/4.0)
|   |   Age >= 33.5
|   |   |   RegDaysEarly < 10.5: T(33.0/26.0)
|   |   |   RegDaysEarly >= 10.5: T(125.0/37.0)
|   NumPosts2 >= 21.5: T(806.0/95.0)
  
```



J48 Classifier Output

Accuracy: .764

Precision: .763

Recall: .765

Future Directions

- Further analyze longitudinal data for discussion participation
- Compare semester over semester data
- Create trend data over multiple years
- Bring in more specific data on discussion to assessment correlation