

AdmitSee Analytics

Mike Yung



mikeyung



yungmsh



yungmsh



Context



The Model

- Can we build a model that predicts a student's chances* of being admitted into college?

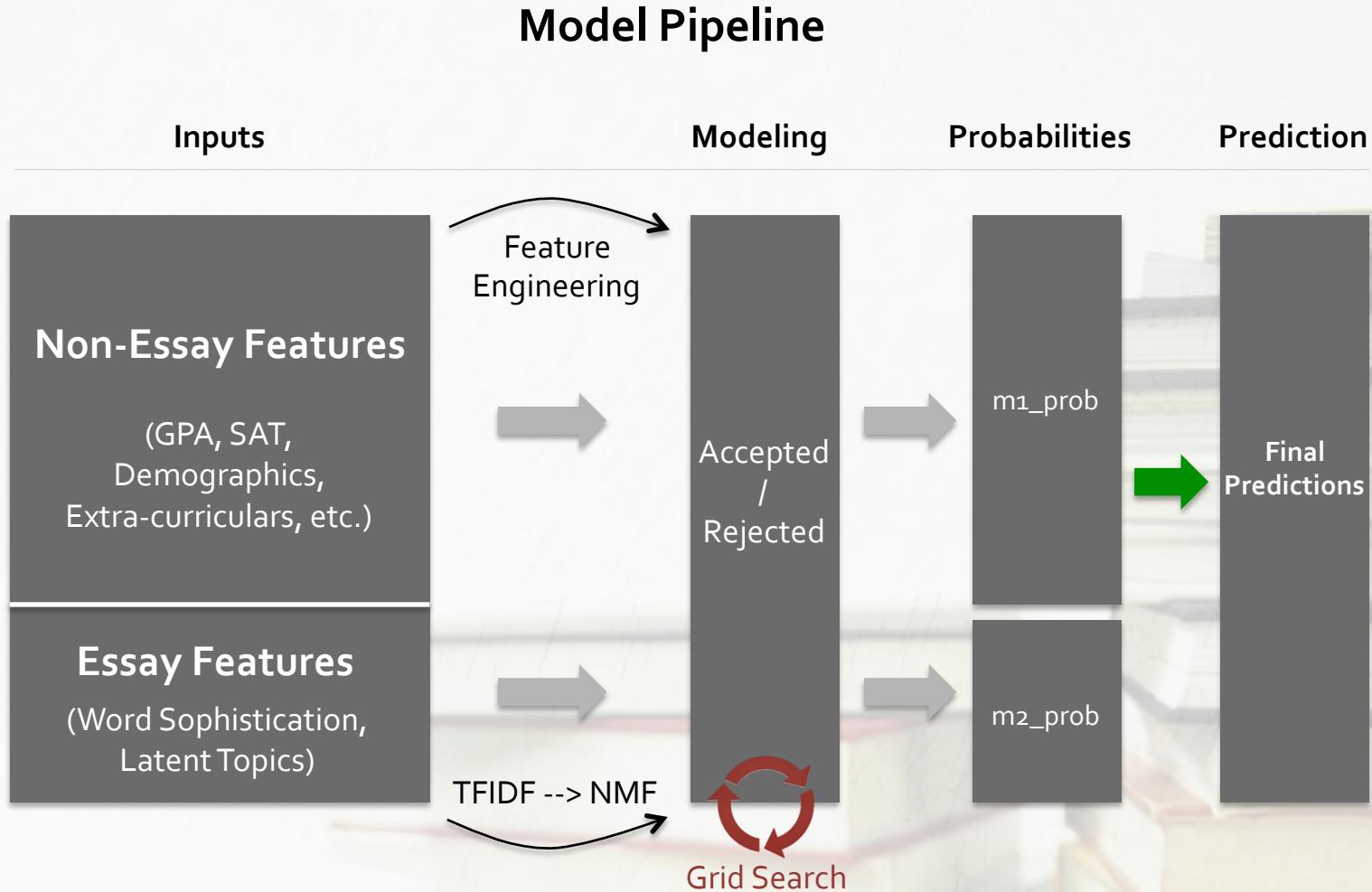


The Essay

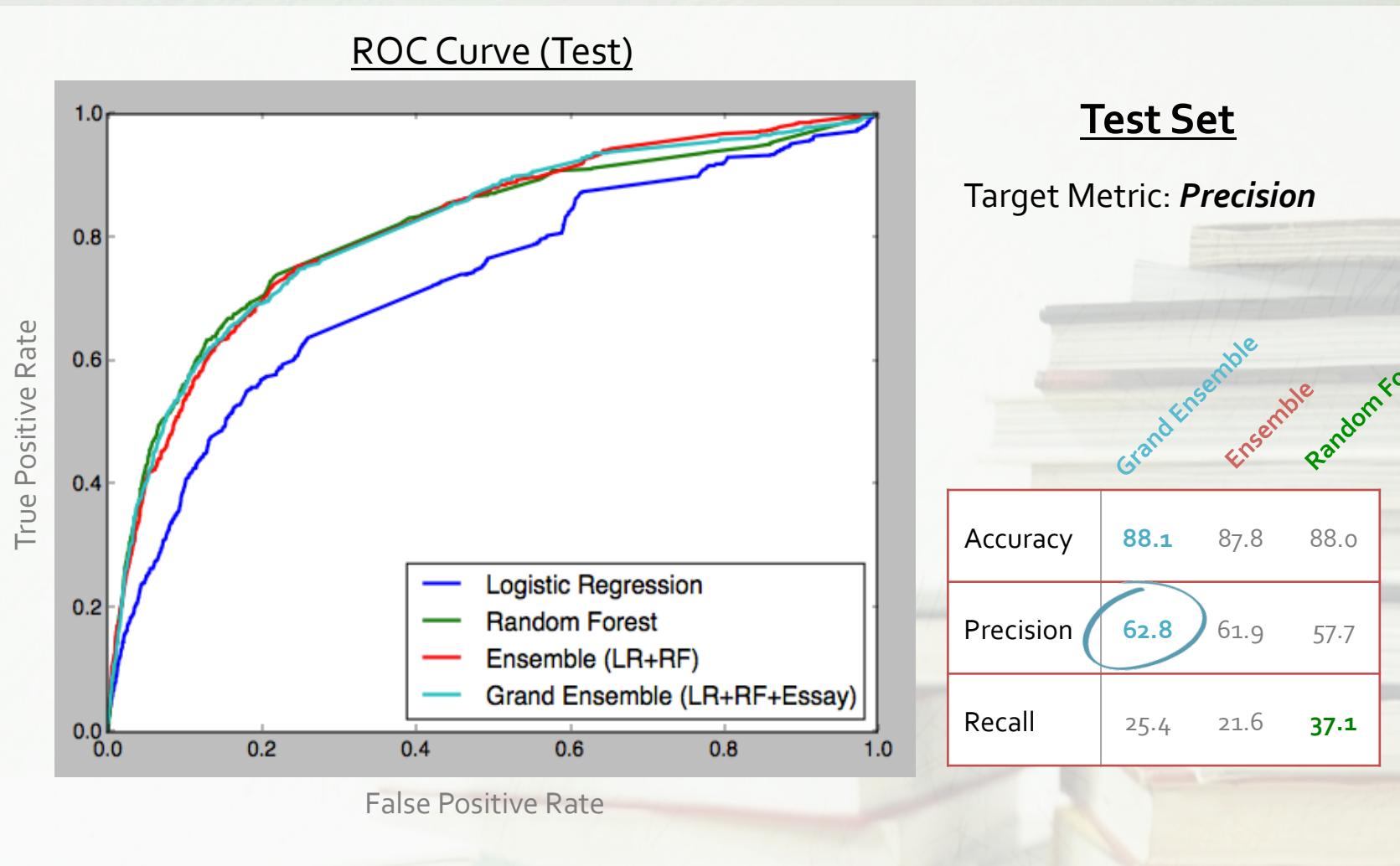
- What insights can we glean from the Common App essay?

*There are some resources that 'calculate' your chances based on your GPA, SAT, and demographics, but none (at least publicly available) that take into account detailed factors such as specific extracurriculars, academic trajectory, the Common App essay etc.

Building an Ensemble Model



Evaluating the Model



Interpreting the Model

Logistic Regression Model

Variable	$e^{\text{Coefficient}}$
Leader	2.26
Student Gov	1.69
Varsity Sport	1.58
Sports Captain	1.29
Award	1.22
Community Service	1.21
SAT Score	1.0003
SAT Times Taken	0.39

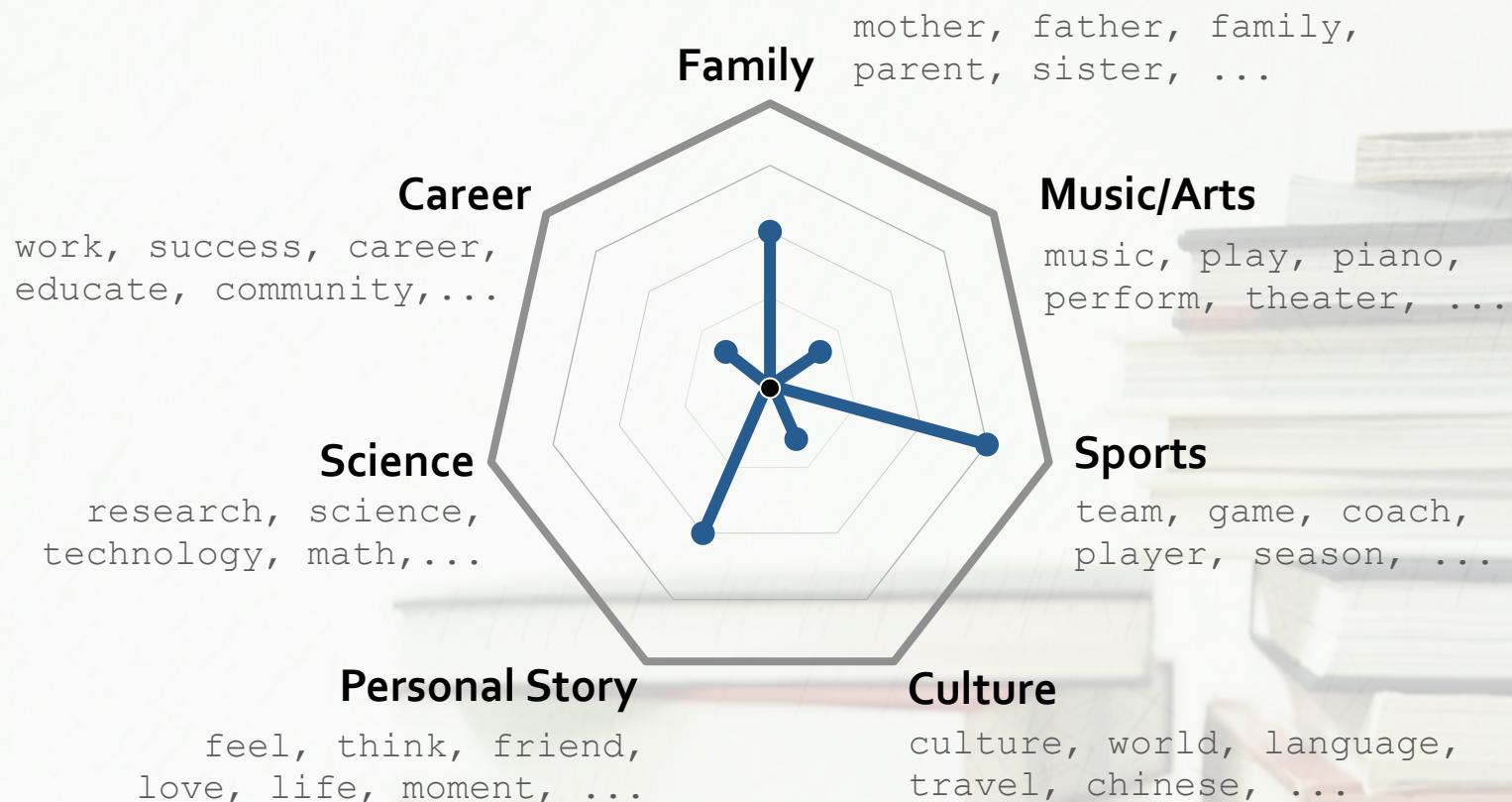
How to Interpret?

If you aren't already in a **leadership** position, taking one will **more than double** your odds of being admitted

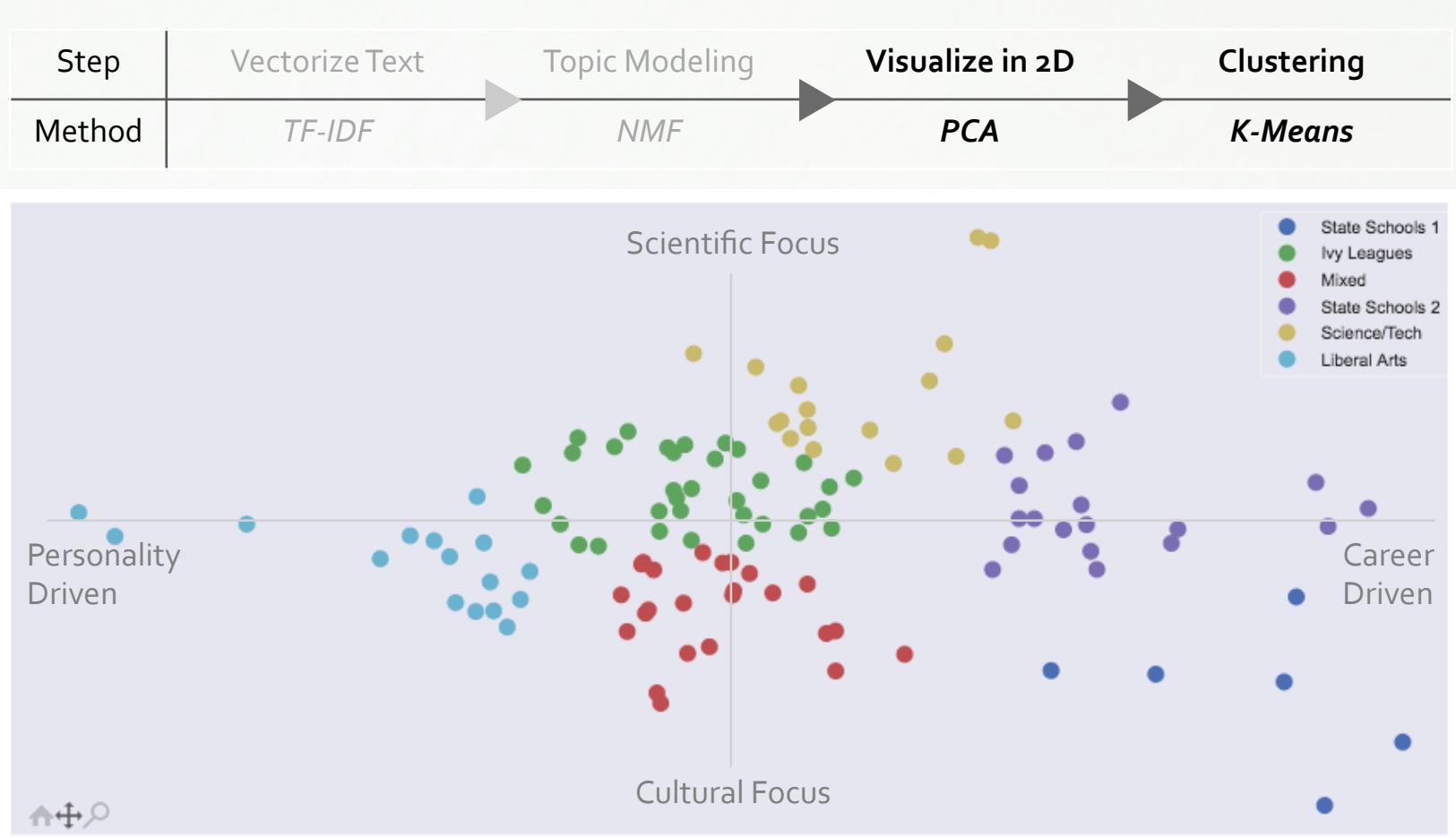
*Note: this is only a subset of all variables used in the model

What's in an Essay?

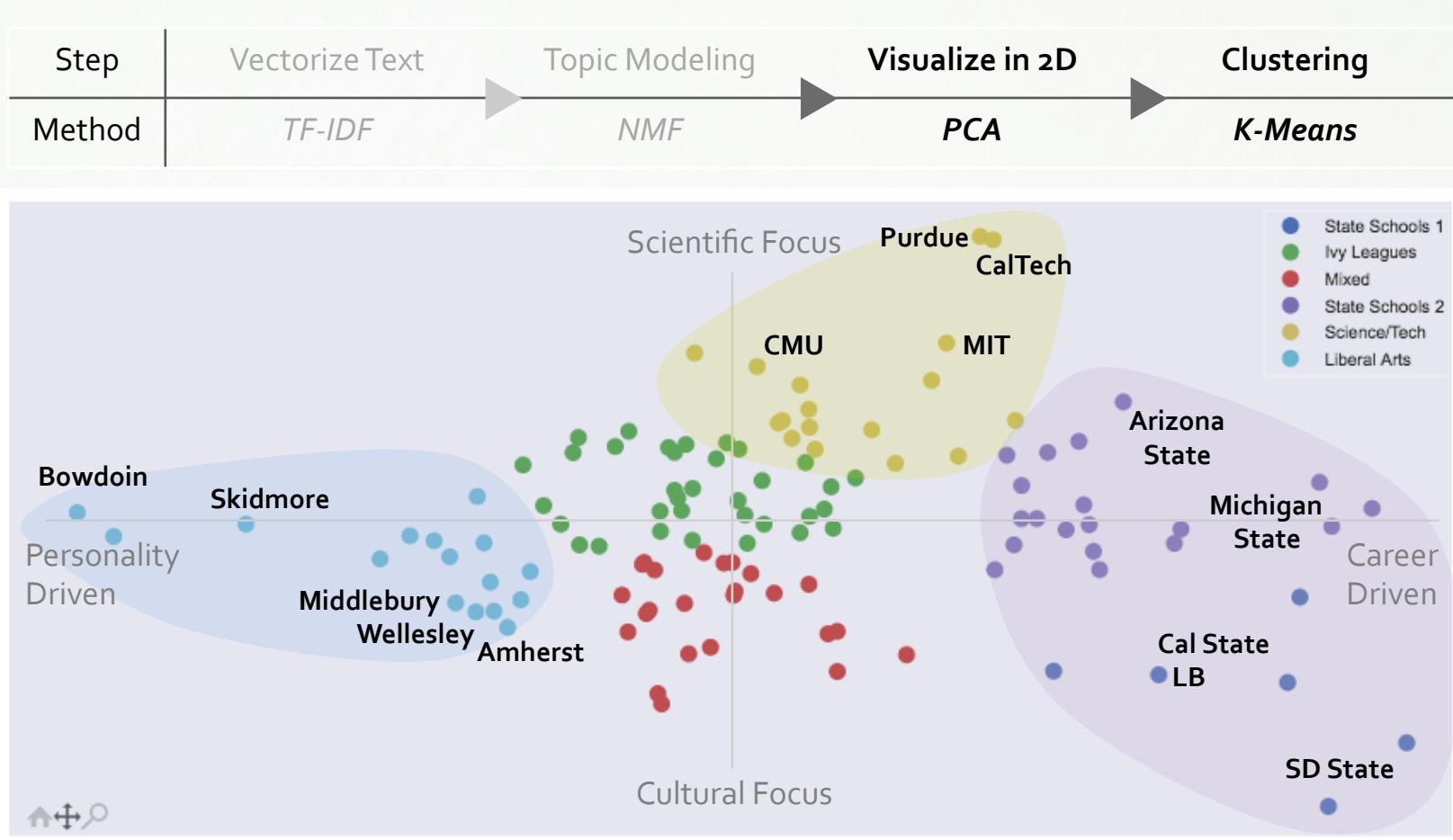
Topic Distribution of a Sample Essay



A 2-D Representation of College Essays

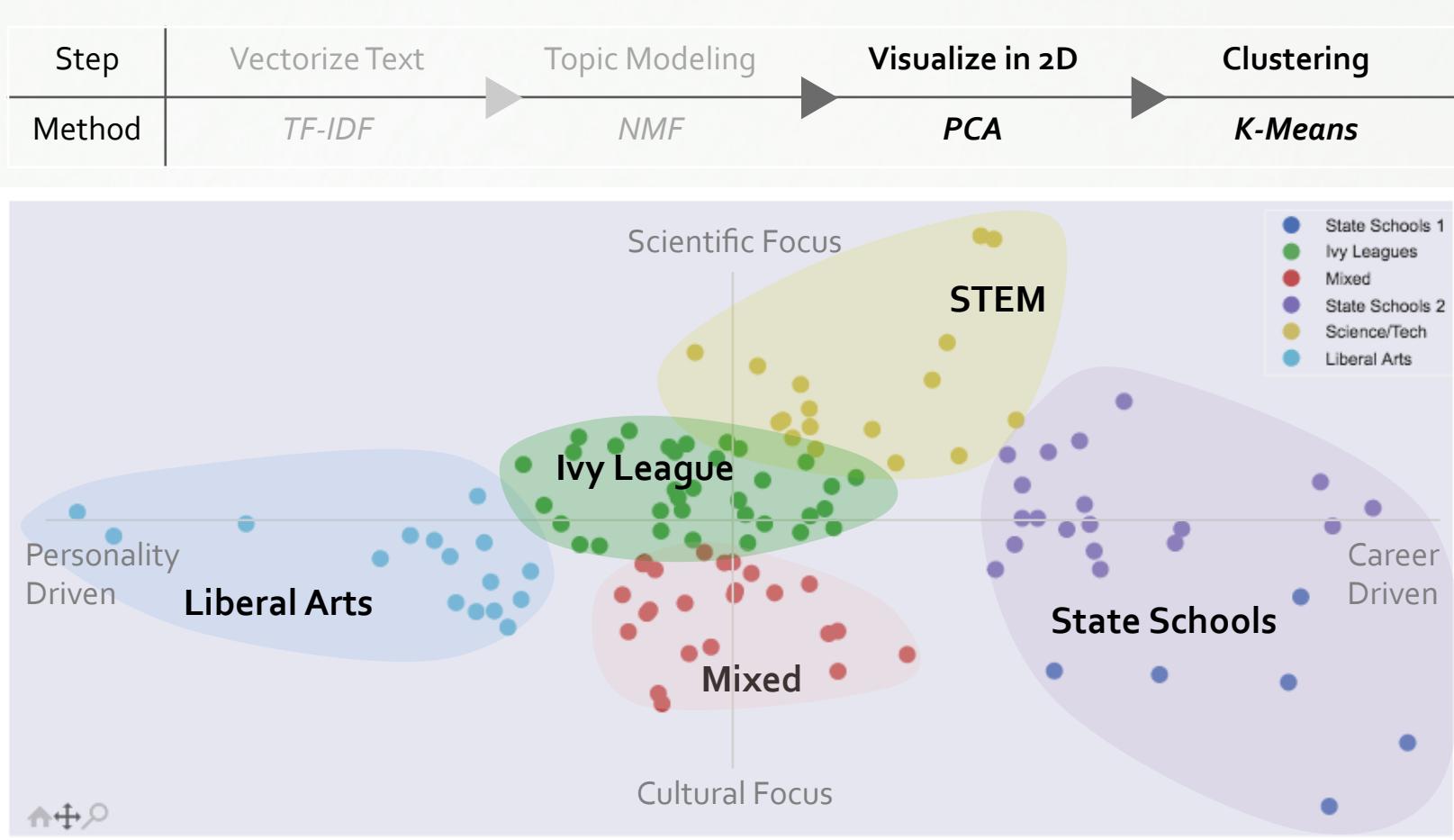


A 2-D Representation of College Essays



*Each point represents a college/university

A 2-D Representation of College Essays



Final Thoughts

- Limitation of data
 - Only enough to model 'top school' admittance
 - With more data:
 - School-level model
 - Graduate school model
- Further NLP exploration
 - N-gram (bigrams, trigrams, etc.)
 - PoS tagging
 - Latent Dirichlet Allocation (LDA)



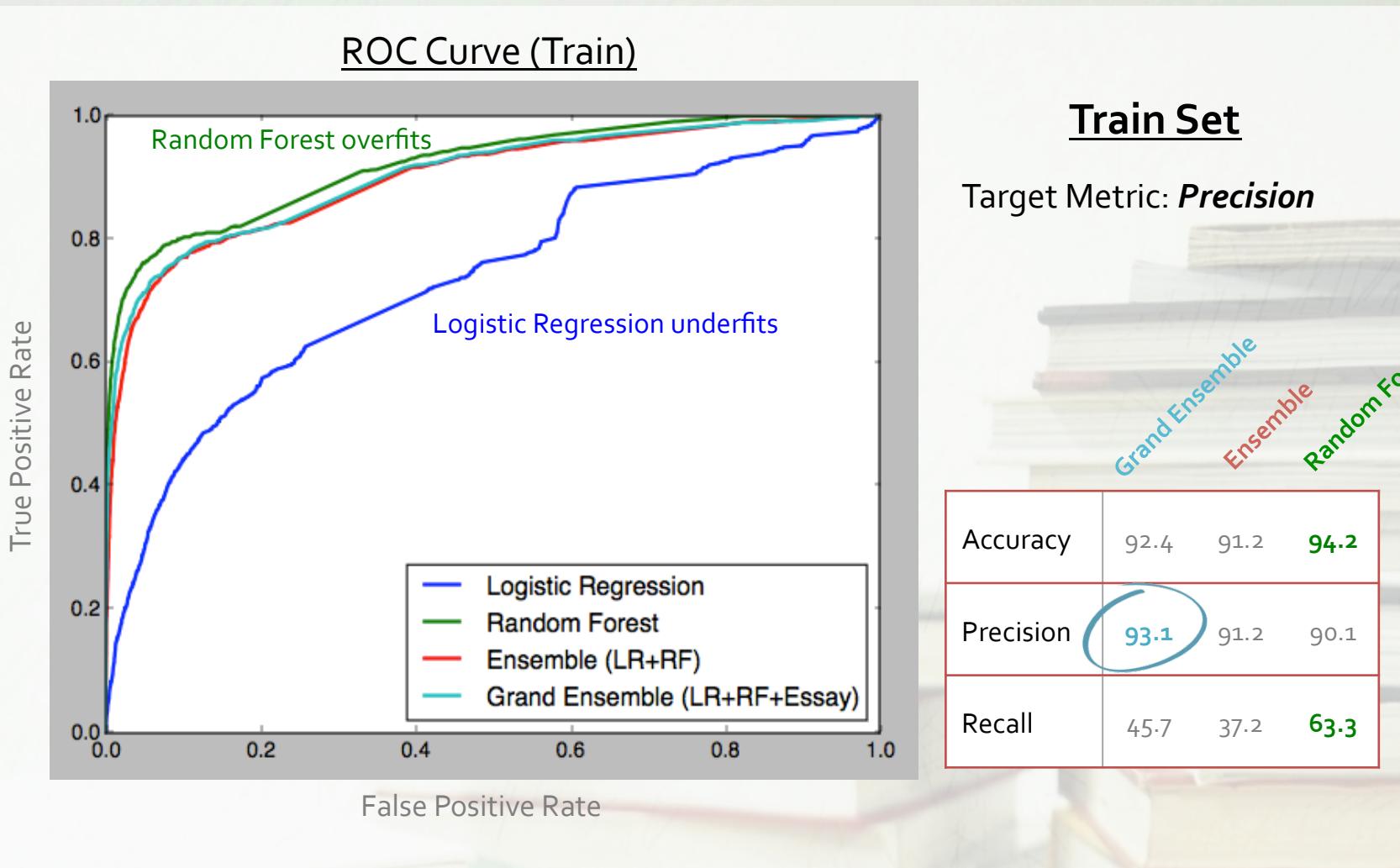
Did You Know?

- If you gain an **additional 100 points** on your **SAT**, you can **increase** your odds of being admitted by **3%**
- If you **take the SATs one more time**, you can **reduce** your odds of being admitted by **61%**

Thank You



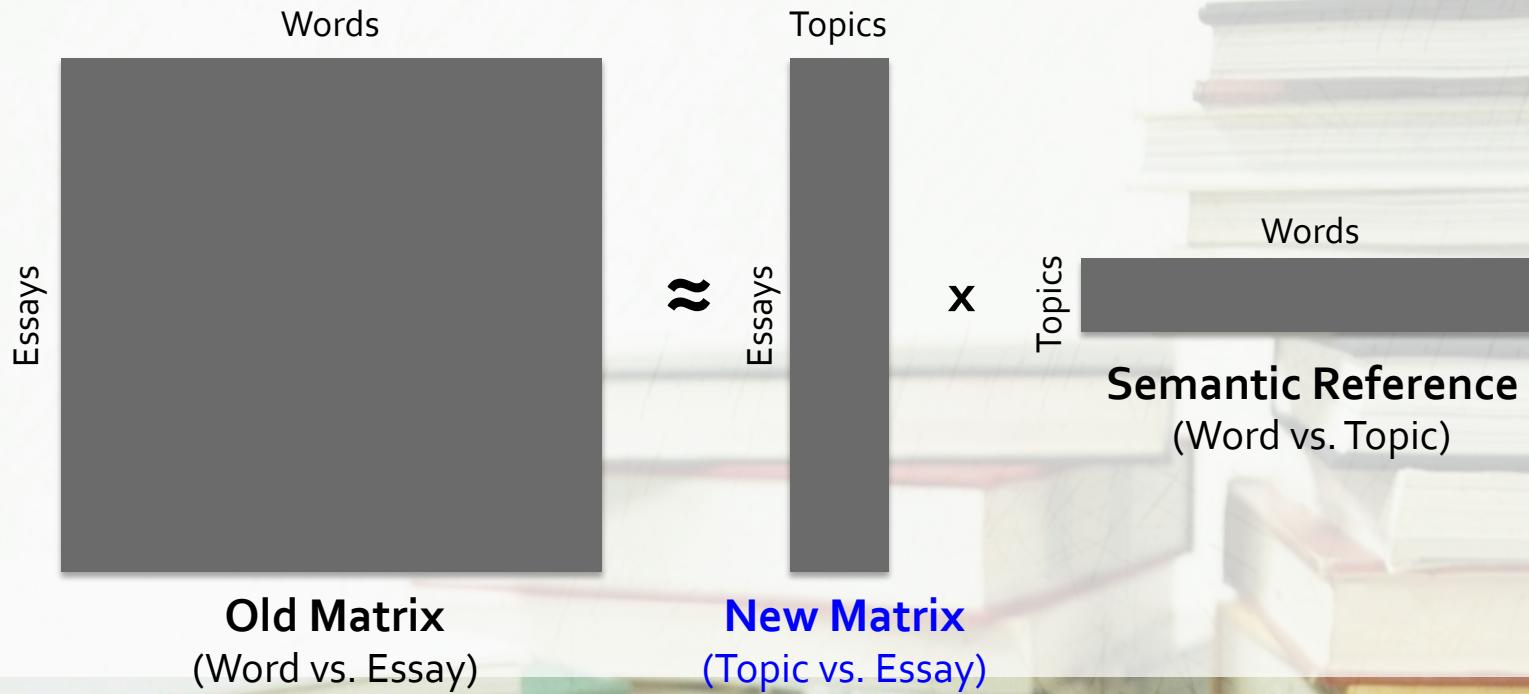
Appendix 1: ROC Curve for Train Set



Appendix 2: NMF Visualized



Visualization of NMF



Appendix 3: NMF Semantic Reference Table



Non-Negative Matrix Factorization

Words

Topics

mother	father	family	parent	sister	...	Family
music	play	piano	perform	theater	...	Music/Arts
culture	world	language	travel	american	...	Culture
team	game	coach	player	season	...	Sports
research	science	computer	technology	math	...	Personal/Story
work	education	career	success	community	...	Science
						Career

Semantic Reference (Word vs. Topic)