Towards Comprehensible Predictive Modeling

Michael Gleicher

Department of Computer Sciences

University of Wisconsin Madison

What is a <u>Good</u> model?

What is a <u>Good</u> model?

Accuracy

What is a <u>Good</u> model?

Descriptive Accuracy

Predictive Accuracy

What is a <u>Good</u> model?

Descriptive Accuracy

Predictive Accuracy

Generalizability

Efficiency in use

Efficiency in construction

Robustness

Data required to build

Conciseness

Compactness

Pre-scriptive power

Self-consistency

Validatability

Verifyability

Simplcity

Parsimony

Tradeoffs

What is a Good model?

Descriptive Accuracy
Predictive Accuracy
Generalizability
Efficiency in use
Efficiency in construction
Robustness
Data required to build
Conciseness
Compactness

Pre-scriptive power
Self-consistency
Validatability
Verifyability
Simplcity
Parsimony

Comprehensibility

What is Comprehensibility?

The ability of **some one**
to understand **some thing**
for **some reason**

This is independent of...
where application, model type,
or **how** we help them do it

Who, What, Why

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developers	Inputs	Improve Performance
Data Scientists	Methods	Build Theory
Domain Experts	Model	Extend/Characterize
Audience	Outputs	Build Trust
	Experiments	Actionability

Who, What, Why

Who?

Stakeholder

Develope

Data Sci

Domain

Audience

What?

Outputs

Experiments

Why?

ing

rformance

racterize

Build Trust

Actionability

Modeling is a process

There are many people involved

We should try to help all of them

Who, What, Why

Who?

Stakeholders

Developers – general purpose methods

Data Scientists – general purpose process

Domain Experts – specific applications

Audience – uses the results

What?

Phase of the **Process**

Experiments

Why?

Reason for Wanting

ove Performance

l Theory

nd/Characterize

l Trust

Actionability

Who, What, Why

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developes	Inputs	Improve Performance
Data Scientists	Methods	Build Theory
Domain Experts	Model	Extend/Characterize
Audience	Outputs	Build Trust
	Experiments	Actionability

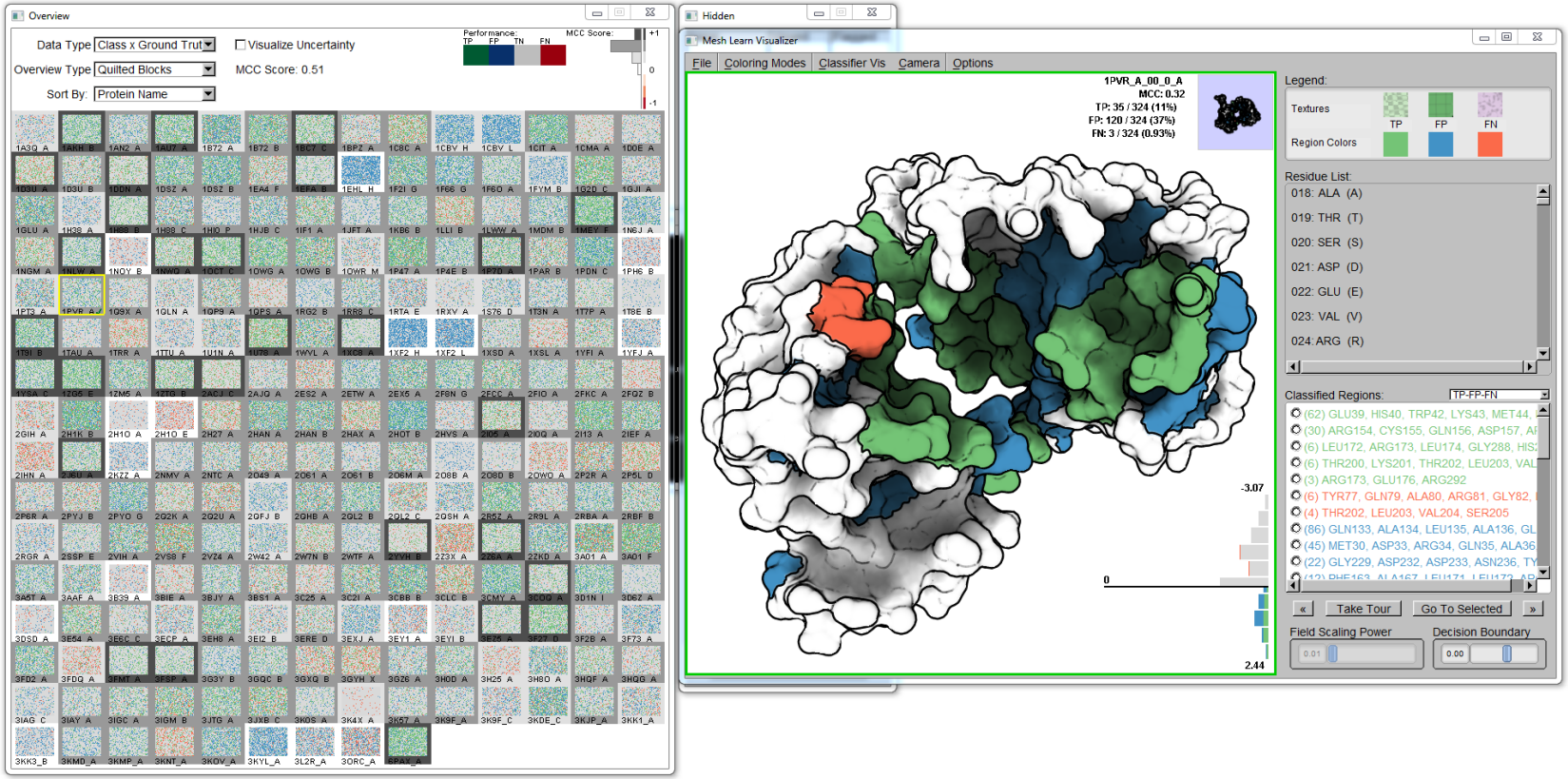
Who, What, Why

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developes	Inputs	Improve Performance
Data Scientists	Methods	Build Theory
Domain Experts	Model	Extend/Characterize
Audience	Outputs	Build Trust
	Experiments	Actionability

Mix and Match

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developers	Inputs	Improve Performance
Data Scientists	Methods	Build Theory
Domain Experts	Model	Extend/Characterize
Audience	Outputs	Build Trust
	Experiments	Actionability

Visual exploration of modeling validation experiments



Sarikaya, Albers, Mitchel and Gleicher. *Visualizing Validation of Protein Surface Classifiers.* Computer Graphics Forum. Proceedings EuroVis 2014

Our Motivation

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developers	Inputs	**Improve Performance**
Data Scientists	**Methods**	Build Theory
Domain Experts	Model	Extend/Characterize
Audience	Outputs	Build Trust
	Experiments	Actionability

Pleasant Surprises

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developers	Inputs	Improve Performance
Data Scientists	Methods	Build Theory
Domain Experts	Model	Extend/Characterize
Audience	Outputs	**Build Trust**
	Experiments	Actionability

Pleasant Surprises

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developers	Inputs	Improve Performance
Data Scientists	Methods	Build Theory
Domain Experts	Model	Extend/Characterize
Audience	Outputs	**Build Trust**
	Experiments	Actionability

Pleasant Surprises

Who?	What?	Why?
Stakeholders	Phase of the **Process**	Reason for Wanting
Developers	Inputs	Improve Performance
Data Scientists	Methods	**Build Theory**
Domain Experts	Model	Extend/Characterize
Audience	Outputs	Build Trust
	Experiments	Actionability

Agenda

Understand Understandability

Metrics for Measuring

Tools for user control over tradeoffs

- tunable methods
- different methods to give options

3 Paths to Understandability

Easier to understand models

Better explanations/visualizations of existing models

Co-design of models and visualizations

3 Paths to Understandability

Easier to understand models

- Explainers (VAST 2013)

Better explanations/visualizations of existing models

- Protein Surface Classifier Validation (EuroVis14)
- Serendip (VAST14)

Co-design of models and visualizations

- Molecular Motion Illustration (TVCG12)
- Splatterplots (TVCG13)

Thanks!

To you for listening.
To the organizers for inviting me
To my students and collaborators.
To the NSF and Mellon Foundation for funding.

Towards Comprehensible (Predictive) Modeling

Michael Gleicher

University of Wisconsin Madison

`gleicher@cs.wisc.edu`