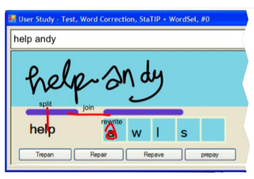
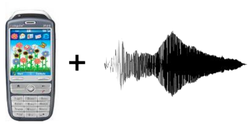
Examining Obstacles to Software Developer Adoption of Statistical Machine Learning

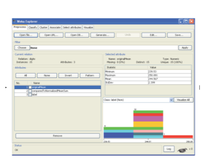
Machine Learning Is Important





Many pieces of software use machine learning algorithms to solve complex tasks

But Existing Tools Are Hard To Use



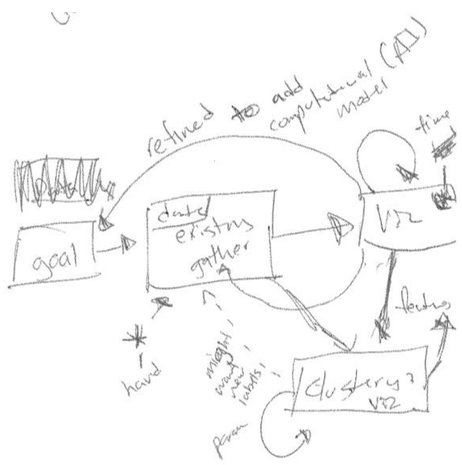
78%

WEKA

Tools designed with a focus on comparing algorithms using benchmark data, not the needs of developers applying machine learning

Two Studies of Developer Needs

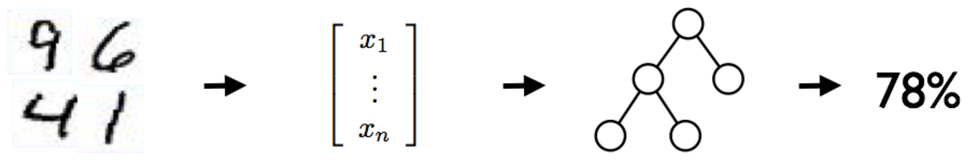
Semi-Structured Interviews



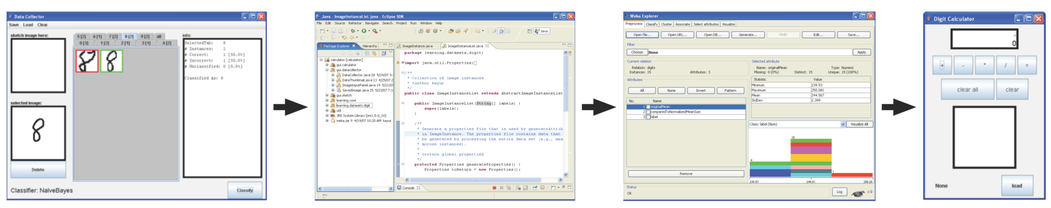
11 researchers recall, discuss, and diagram their process of applying statistical machine learning in a recent project

Laboratory Study of Digits Task

10 participants spend 5 hours building and testing a complete learning system

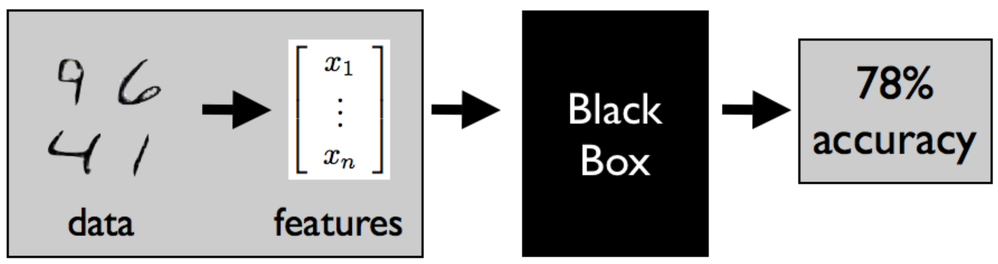


Data Collection → Feature Generation → Algorithm Selection → Testing



Three Obstacles

Difficulty *understanding* the relationship between data and algorithm output



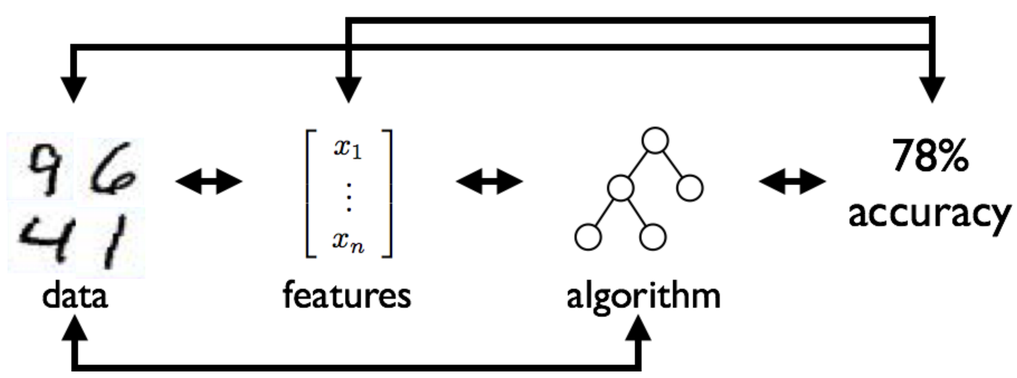
"If you have a black box that is a [statistical machine learning algorithm] and it produces numbers in the end, then you have no idea what actually happened."

Difficulty *evaluating* algorithms within the context of applications

"the cross-validation would show ... 85% to 90% accuracy ... and then you would try it ... it worked extremely well for some people and not well for others.

Difficulty effectively pursuing an *iterative and exploratory* process

"We basically tried a whole bunch of Weka experimentation and different algorithms ... and nothing worked, so we decided that ... maybe we should explore the feature space."



Investigating Statistical Machine Learning as a Tool for Software Development. *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI 2008).

Examining Difficulties Software Developers Encounter in the Adoption of Statistical Machine Learning. *Proceedings of AAAI Conference on Artificial Intelligence* (AAAI 2008).

Kayur Patel, James Fogarty, James Landay, Beverly Harrison

design: use: build: university of washington