

Survey of Machine Learning Methods and Resources

John D. Lee and Linda Ng Boyle

10/1/2018

Machine Learning Next Steps

- Machine learning challenges and opportunities
- Other machine learning techniques
- Machine learning resources

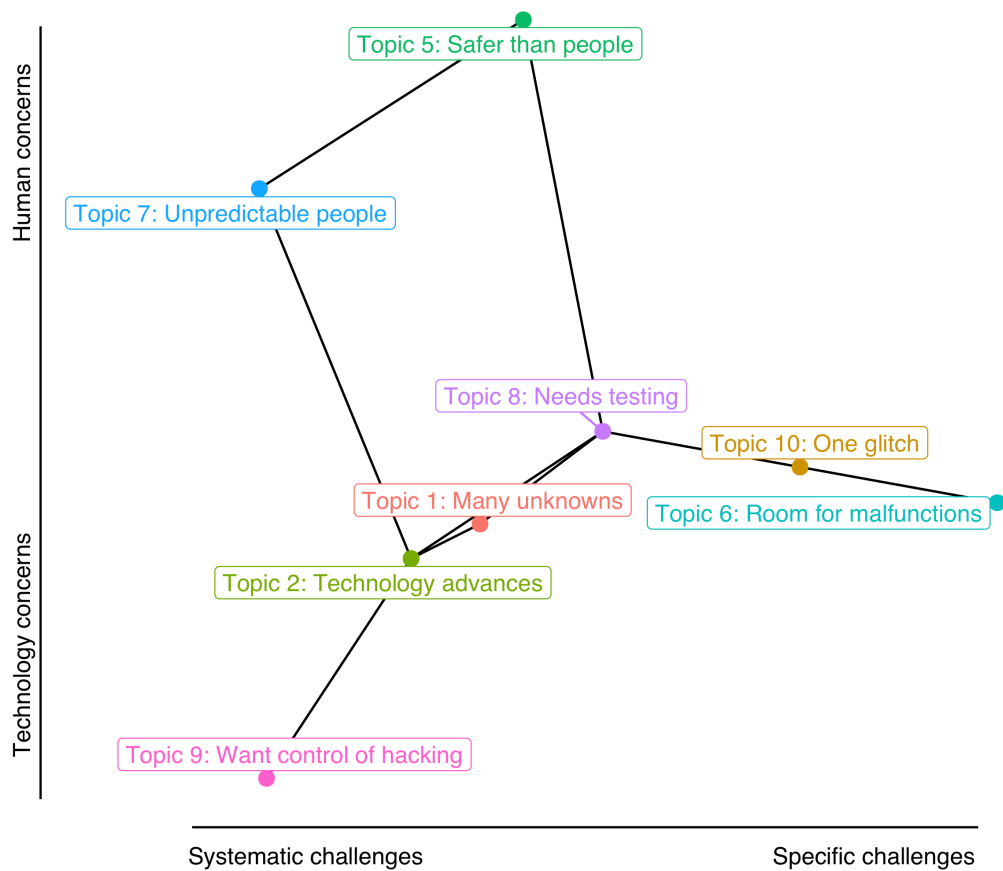
Machine Learning Challenges and Opportunities

Issues that behavioral scientists can help address

- Representative sampling
- Differential costs of prediction errors
- Tradeoff of model understandability and performance
- Explaining model output (LIME)
- Ethical considerations
- Design thinking and systems thinking

Machine Learning Techniques

Text analysis: Quantitative analysis of qualitative data



Machine learning resources: caret

- Comprehensive online description of caret features (<http://topepo.github.io/caret/index.html>)
- Early draft of book on feature engineering by the author of caret (<http://www.feate.engineering>)

Machine Learning Resources: Keras

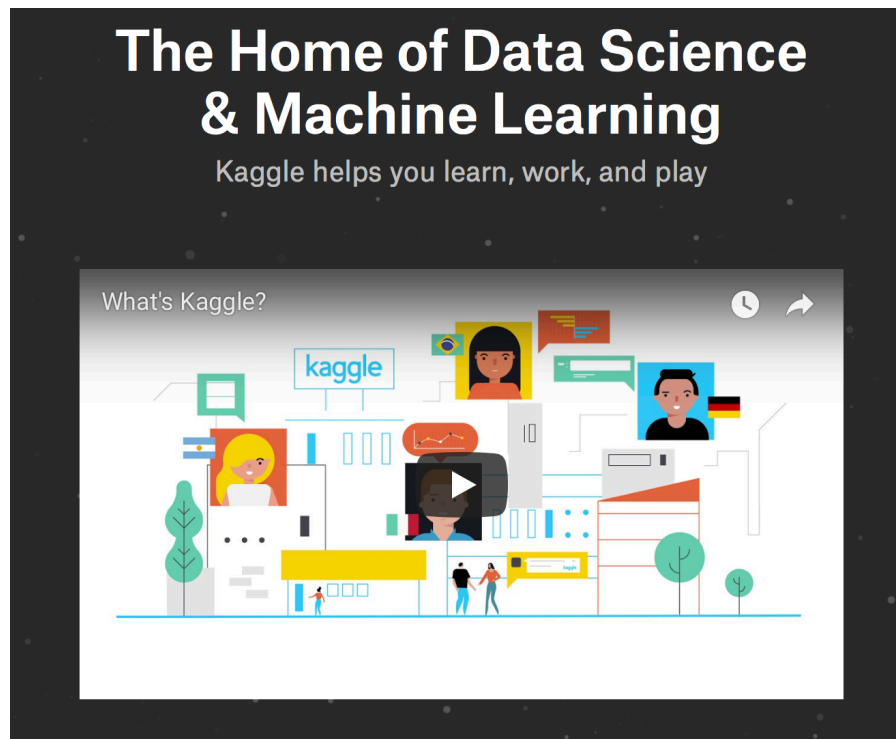
- Simple interface to powerful deep learning algorithms
- Support for prototyping through deployment



<https://keras.rstudio.com>

Machine Learning Resources: Kaggle

- Data—over 10,000 datasets
- Competitions
- Kernels—code developed for analysis of datasets hosted on kaggle



Machine Learning Resources: DataCamp

- Online courses with integrated exercises
- Including basic R and advanced courses
- Machine learning toolbox by the author of caret
- Writing functions by author of the tidyverse



<https://www.datacamp.com/courses/machine-learning-toolbox>

Workshop Summary

Human Factors contributions to ML and ML contributions to Human Factors

- Data visualization: Don't venture into uncharted territory
- Data reduction and cleaning: 80% of any data analysis
- Multi-level general linear models
- Machine learning overview and implementation
- Survey of machine learning behavioral issues and applications

Repeatable data analysis with R and RStudio