



DATA AT BIG CARTEL

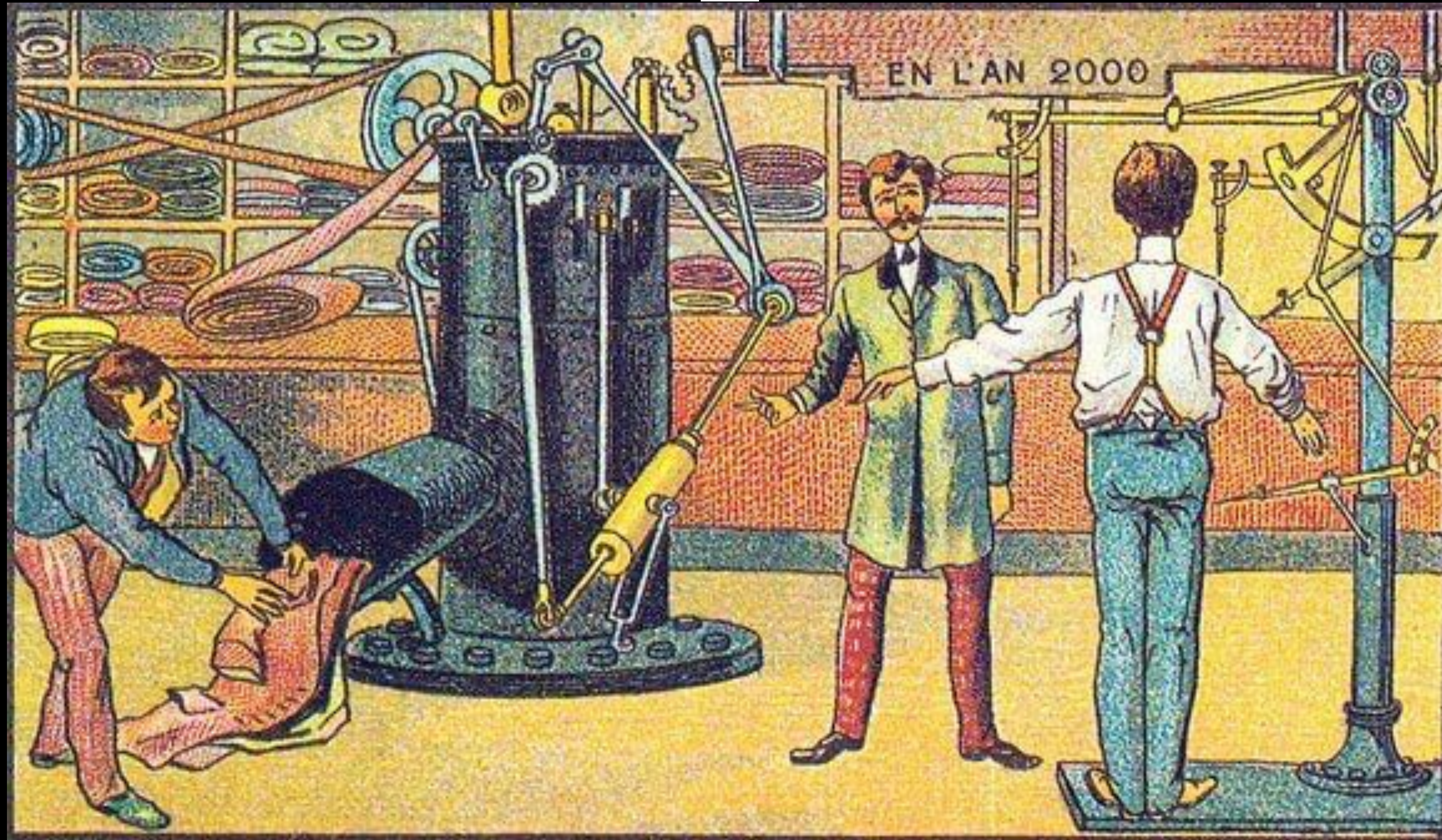
MACHINE LEARNING IS
FOR EVERYONE!

THE PLAN

- What even is “Machine Learning”?
- SciKit Learn Example Pages
- Example from my work, Big Cartel



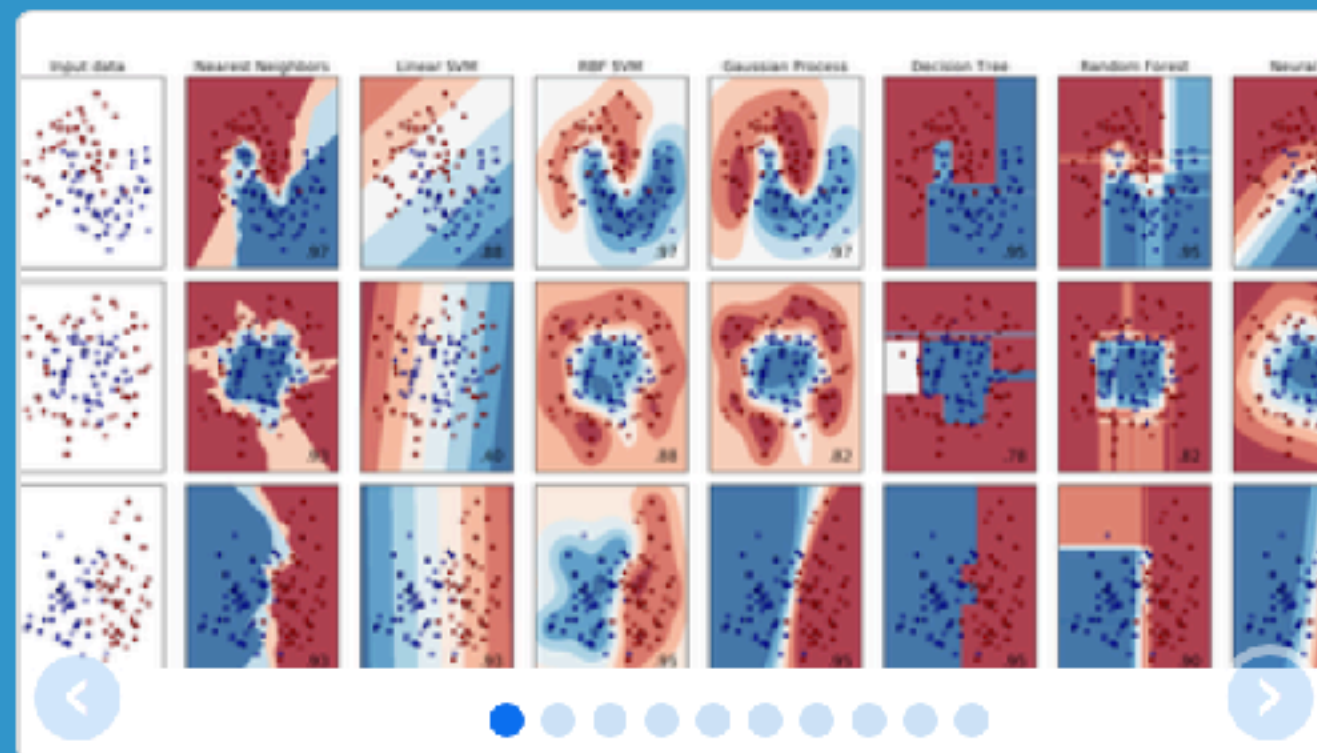
MACHINE LEARNING



<https://singularityhub.com/2012/10/15/19th-century-french-artists-predicted-the-world-of-the-future-in-this-series-of-postcards/>



SCIKIT LEARN



scikit-learn

Machine Learning in Python

- Simple and efficient tools for data mining and data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

Classification

Identifying to which category an object belongs to.

Applications: Spam detection, Image recognition.

Algorithms: SVM, nearest neighbors, random forest, ... — Examples

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices.

Algorithms: SVR, ridge regression, Lasso, ... — Examples

Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes

Algorithms: k-Means, spectral clustering, mean-shift, ... — Examples

Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency

Algorithms: PCA, feature selection, non-negative matrix factorization. — Examples

Model selection

Comparing, validating and choosing parameters and models.

Goal: Improved accuracy via parameter tuning

Modules: grid search, cross validation, metrics. — Examples

Preprocessing

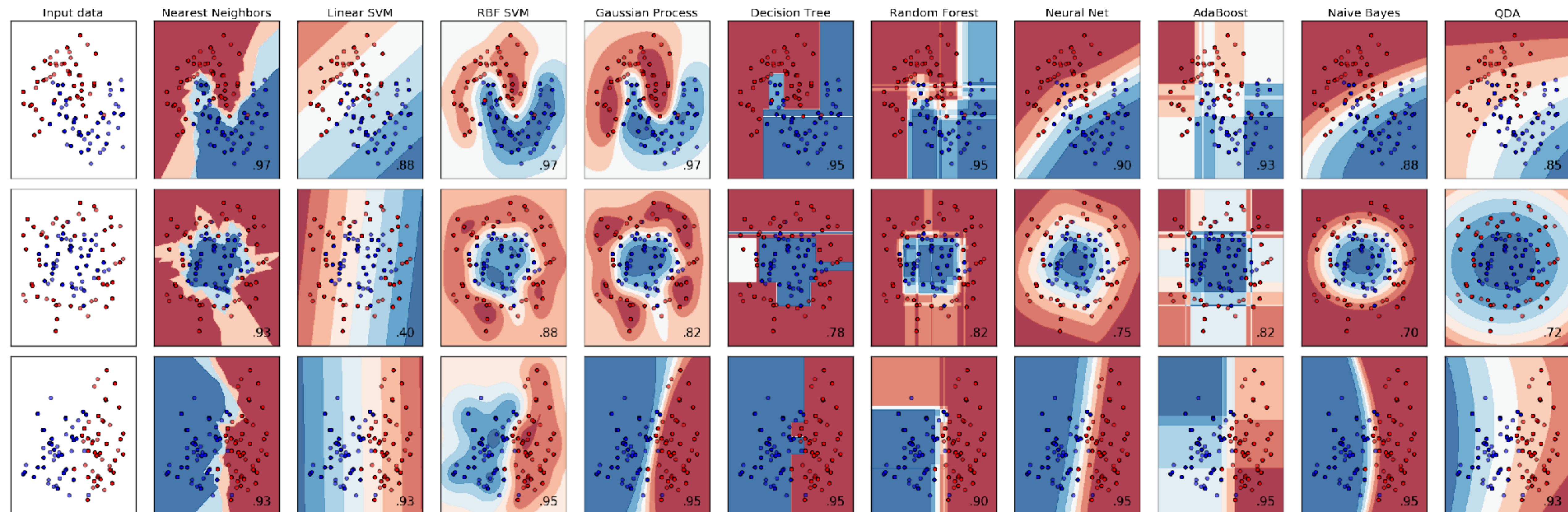
Feature extraction and normalization.

Application: Transforming input data such as text for use with machine learning algorithms.

Modules: preprocessing, feature extraction. — Examples



CLASSIFICATION

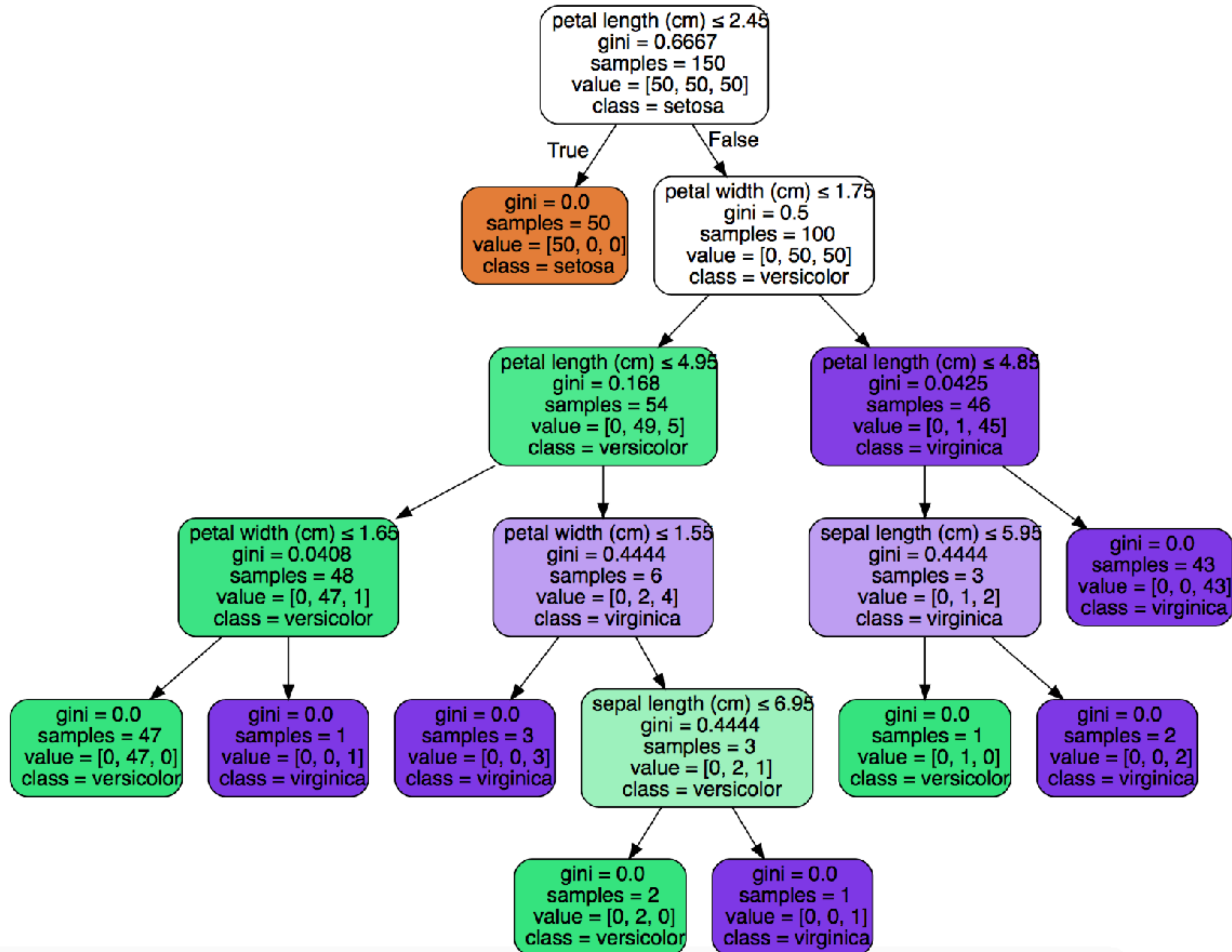


http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html



DECISION TREES

<http://scikit-learn.org/stable/modules/tree.html>



We Believe
in the
Artist



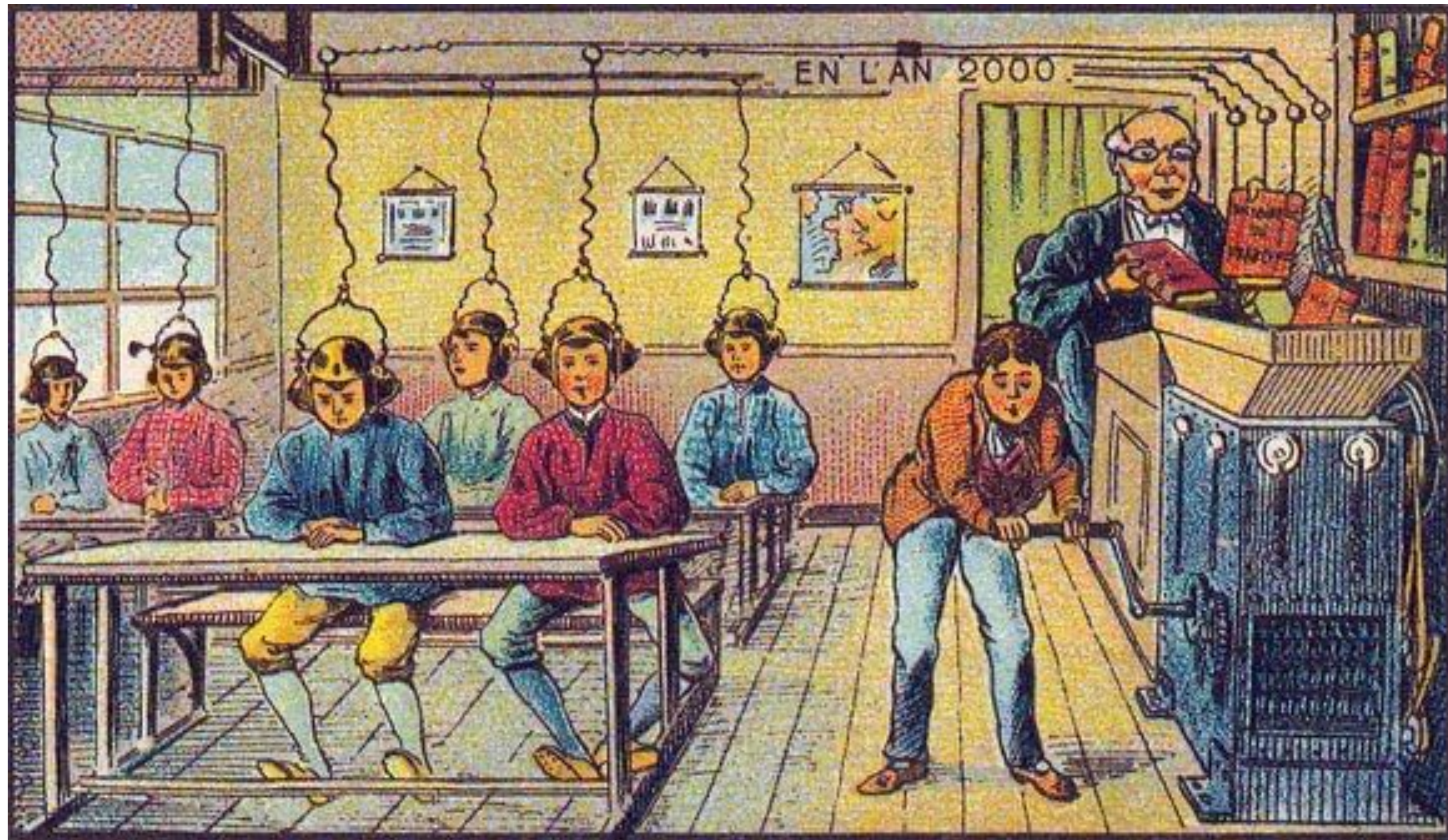
CUT TO JUPYTER NOTEBOOK



SUGGESTED READINGS

- Data Science from Scratch: First Principles with Python, O'Reilly, Joel Grus
- The Elements of Statistical Learning: Data Mining, Inference, and Prediction; Springer; Trevor Hastie, Robert Tibshirani, Jerome Friedman
- Machine Learning: A Probabilistic Perspective, The MIT Press, Kevin P. Murphy
- Pattern Recognition and Machine Learning, Springer, Christopher M. Bishop





<https://singularityhub.com/2012/10/15/19th-century-french-artists-predicted-the-world-of-the-future-in-this-series-of-postcards/>