

Today:

- ▶ Tests
- ▶ Review of course organization
- ▶ “Traditional methods” vs “ML methods” example
- ▶ Machine learning core terms and concepts

## First half of the course:

- ▶ Introduction (Aug 23–25)
- ▶ Regular expressions (Aug 28–30)
- ▶ Edit distance (Sept 1)
- ▶ Information theory (Sept 6–8)
- ▶ Language models (Sept 11–18)
- ▶ Parts of speech and HMMs (Sept 18–27)
- ▶ Parsing (Sept 29–Oct 4)
- ▶ Review (Oct 6–9)
- ▶ Midterm (Oct 11)

## Second half of the course:

- ▶ Introduction (Oct 13–20)
- ▶ Naïve Bayes Classification (Oct 23–27)
- ▶ Stylometry (Oct 30–Nov3)
- ▶ Word embeddings (Nov 6–10)
- ▶ Neural nets (Nov 13–17)
- ▶ Machine translation (Nov 20–27)
- ▶ Text generation (Nov 29–Dec 1)
- ▶ Ethics (Dec 4–6)

*Machine learning is a form of applied statistics with emphasis on the use of computers to statistically estimate complicated functions.*

*Goodfellow et al., Deep Learning, 2016. Pg 95.*

*Machine learning is the science (and art) of programming computers so they can learn from data. [In 1959, Arthur Samuel defined machine learning as the] field of study that gives computers the ability to learn without being explicitly programmed.*

*Géron, Hands-On Machine Learning, 2019. Pg 2.*

*[Machine learning is] a set of methods that can automatically detect patterns in data and then use the uncovered patterns to predict future data or to perform other kinds of decision-making under uncertainty.*

*Murphy, Machine Learning: A Probabilistic Perspective, 2012. Pg 1.*

## Machine learning main tasks:

- ▶ Regression, where the target type is  $\mathbb{R}$
- ▶ Classification, where the target type is a finite set
  - ▶ Binary classification, where the target is  $\{F, T\}$  (or  $\{0, 1\}$  or  $\{-1, 1\}$  ...)
- ▶ Density estimation, where the target type is  $[0, 1]$ .

## Other machine learning tasks (see Goodfellow, *Deep Learning*, pg 96–100):

- ▶ Transcription, where the observations are unstructured and the targets are text.
- ▶ Machine translation, where the observations and targets are text.
- ▶ Anomaly detection, where the targets are indicators of whether the observation is atypical.
- ▶ Synthesis and sampling, where there are no observations in deployment, but rather the program produces new observations similar to those in training.
- ▶ Denoising, where the targets are corrected versions of the observations.

Coming up:

- ▶ Fall break