Predictive Analytics Lecture 1

Adam Kapelner

Stat 422/722
at The Wharton School of the University of Pennsylvania

January 17 & 18, 2017

Define: Prediction and Forecast

“statement about an uncertain event”,

Define: Prediction and Forecast

“statement about an uncertain event”, “informed guess or opinion”

predict (v.) 1620s (implied in predicted), "*foretell, prophesy*," a back formation from prediction or else from Latin praedicatus, past participle of praedicere "foretell, advise, give notice,"

forecast (n.) early 15c., "*forethought, prudence*," probably from forecast (v.). Meaning "conjectured estimate of a future course" is from 1670s.

Examples

We do this all the time, sometimes unknowingly:

“Apple stock will go up tomorrow”, “This condo will sell for $500K”, “Going skiing this weekend will make me happy”,

How?

We use a *model*.

Define: model

"a functional decription of a system" and an example model is:

> *Absolute power corrupts absolutely*

All models have **input(s)** and **output(s)**. What are the input(s) and output(s) above?

Input: Power? Corruption? Likely: a person's amount of power.
Output: Person? Power? Corruption: Likely: that person's amount of corruption.

The model can be rephrased less poetically as:

> *If a person has a lot of power then that person will be very corrupt.*

Define: Prediction

Here, the input and output are *features* or *attributes* or *chracteristics* of a person.

Generally, inputs and outputs are measurements on the *unit of analysis* or *the observation* or *the subject*. Note: I will be using "observation" or the name of the unit itself: "person", "car", etc.

Now we can define prediction a bit better. For a new / heretofore unseen / future / unknown observation, unit (here, person),

$$\underbrace{\text{prediction}}_{\substack{\text{the guessed}\\ \text{output}\\ \text{measurement}}} = \text{model}(\underbrace{\text{observation}}_{\substack{\text{the measured}\\ \text{inputs}}})$$

The above emphasizes the functional relationship of the observation and prediction through the model. Here,

$$\substack{\text{a guessed}\\ \text{amount of}\\ \text{corruption for Bob}} = \text{model}\left(\substack{\text{Bob's measured}\\ \text{amount of power}}\right)$$

Mathematical Modeling

Our model:

> *If a person has a lot of power then that person will be very corrupt*

Extremely imprecise, vague, ill-defined.