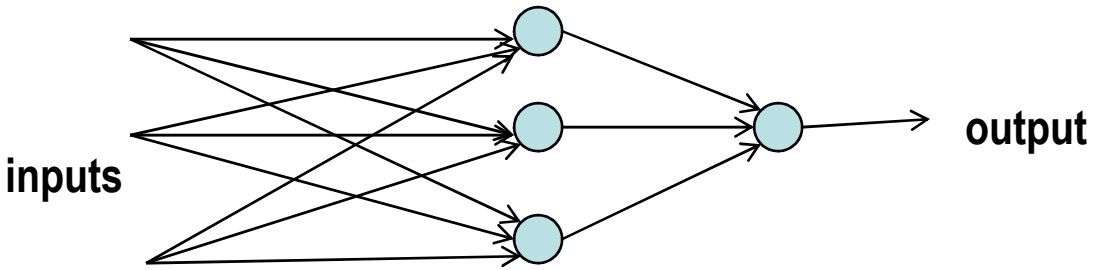
An Overview of Time Series Forecasting for Hotel Revenue Management

Amir Atiya
Dept Computer Engineering
Cairo University
amir@alumni.caltech.edu

Neural Networks

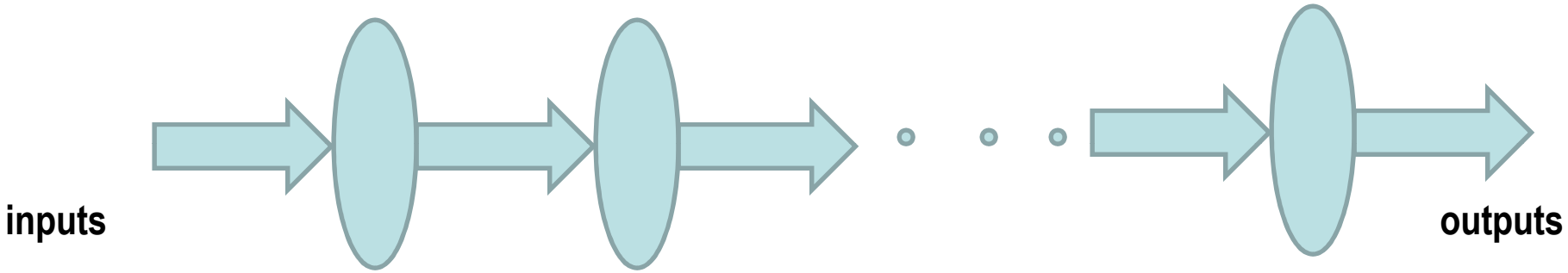
- Typically nonlinear models that learn relation between inputs and outputs, using data driven approaches, or certain probability models.

- Example: Neural networks:
- Networks of “neurons” inspired by the brain’s information processing ability:



Deep Neural Networks

- A NN with many layers

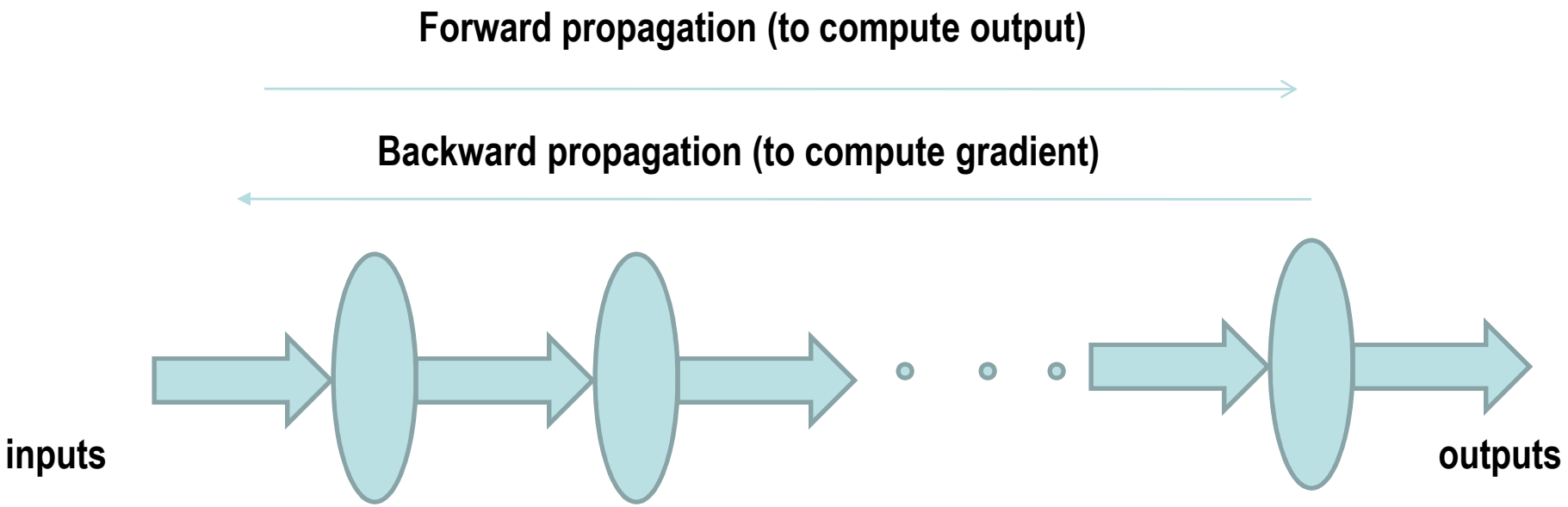


Deep Neural Networks (Contd)

- More flexible than traditional one-hidden-layer network.
- Able to extract useful features from raw data.
- Has proved itself in problems with large number of features.
- Has obtained very good results in computer vision problems.

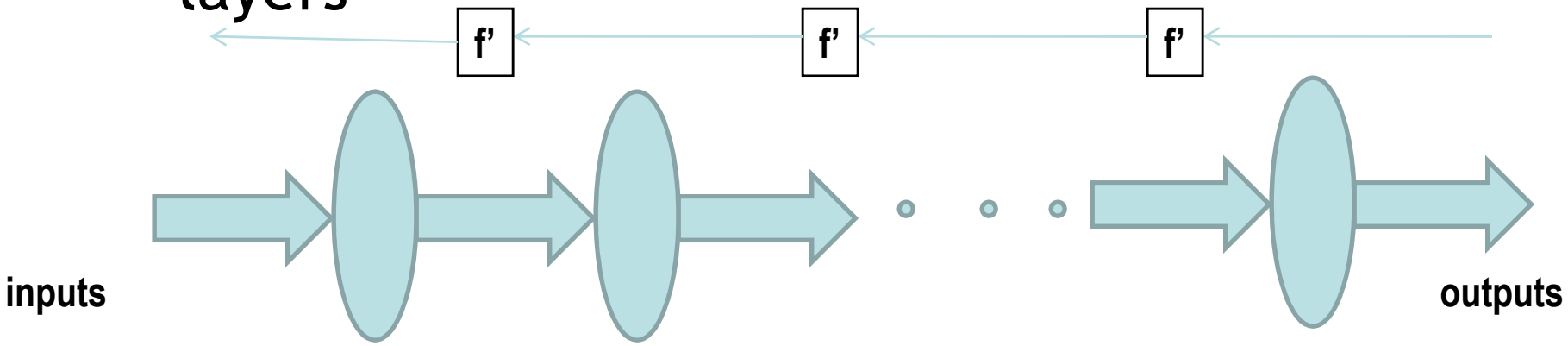
Deep Neural Networks

- Training:



Difficulties of Deep Networks

- Needs a very large data set (to overcome overfitting).
- Need considerable computing resources.
- Training a deep network is harder than shallow networks.
- The reason is the vanishing gradients in deep layers

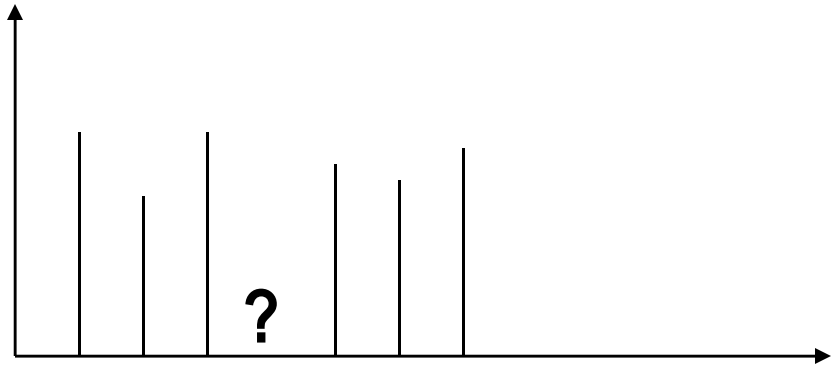


Why Does Deep Networks Work Well

- First layers produce lower level features, and further layers produce progressively more sophisticated features and intermediate variables.
- Some problems are heavily nonlinear.
  - Every layer produces a mildly nonlinear mapping.
  - The concatenation of these mildly nonlinear mappings produce an overall heavily nonlinear mapping.

Time Series Preprocessing

Missing values:



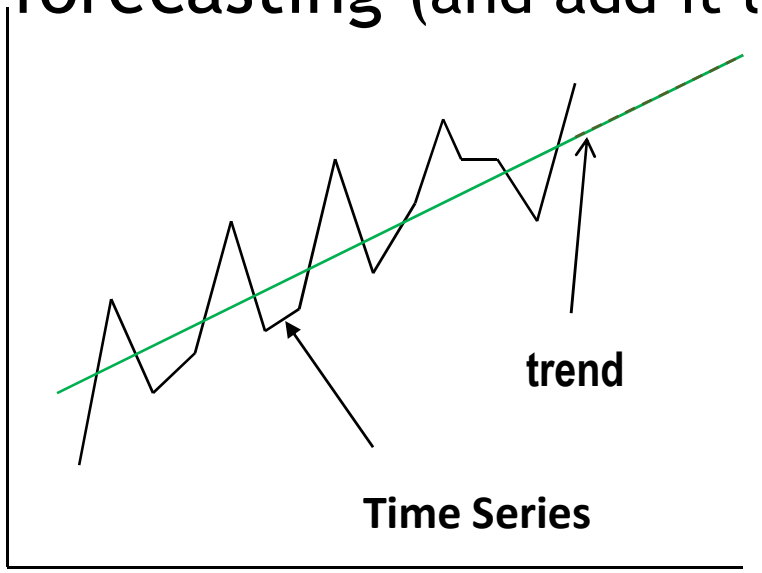
Best approach: Take the average of the preceding and the successive sample.

Another approach is iterative, based on the EM concept.

Time Series Preprocessing

Detrending:

- Remove (subtract) linear trend before applying the forecasting (and add it later).

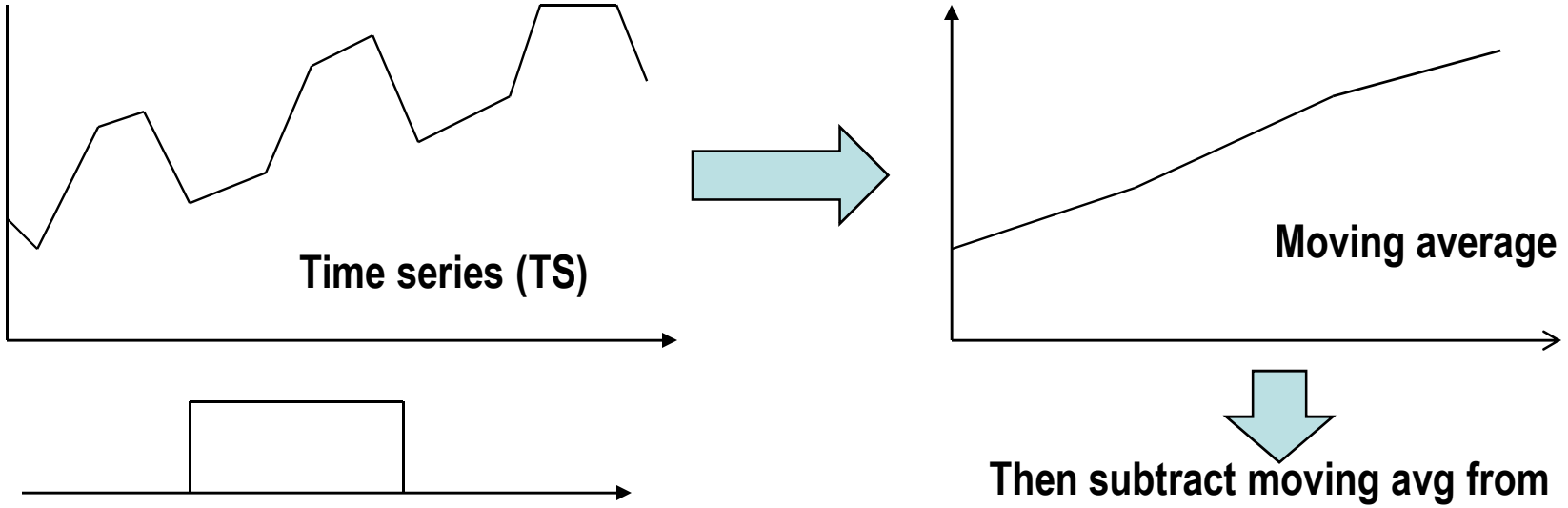


- Nonlinear transformation: e.g. apply **log** transformation on all time series before forecasting.

Time Series Preprocessing

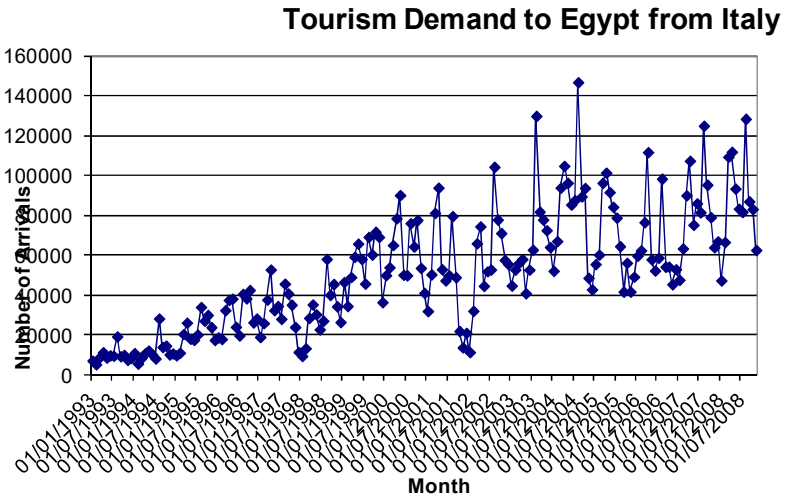
Deseasonalization:

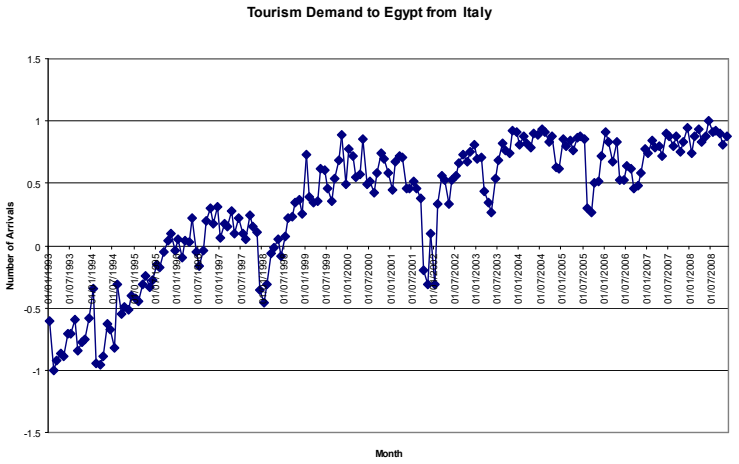
- It is better to remove seasonal or cyclic components before forecasting (and add it later).
- Additive seasonal decomposition method:



Deseasonalization

- Example of a TS (above), and its deseasonalization (below).
- Also included is application of a log transformation.





Seasonal Analysis

- In hotels there are two types of seasonality:
  - Day of the week seasonality: Weekend travel is different from weekday travel.
  - Yearly seasonality: Demand in seasons (summer/winter) is different.

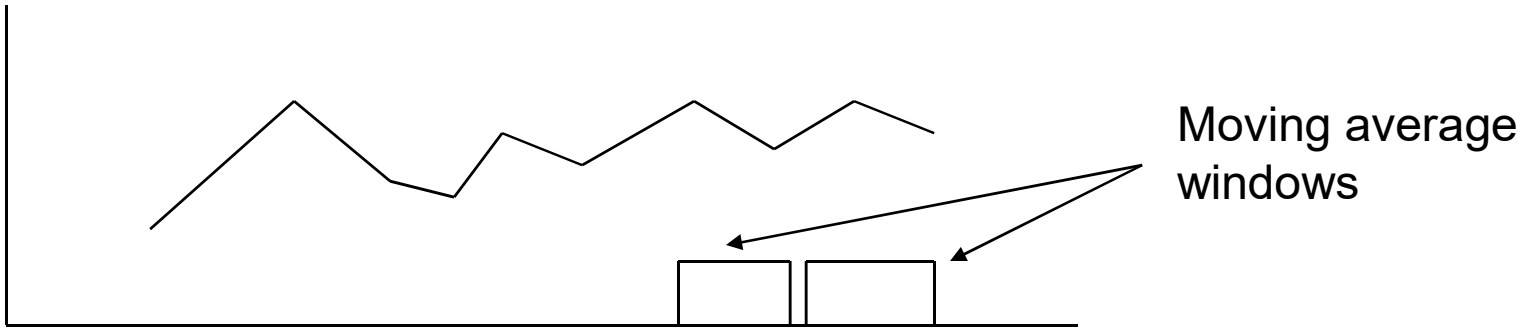
Input Variables

- Past lags $x_{t-m+1}, \ldots, x_t$.
- Differenced TS values:

  $x_{t-m+2}-x_{t-m+1}, \ldots, x_t-x_{t-1}$.
- Moving averages:

$\text{Avg}(x_{t-m+1:t-m+k}), \text{Avg}(x_{t-m+k+1:t-m+2k}), \ldots, \text{Avg}(x_{t-k+1:t})$.

- Best ways: past lags and moving averages.



Multi-Step ahead Forecasting

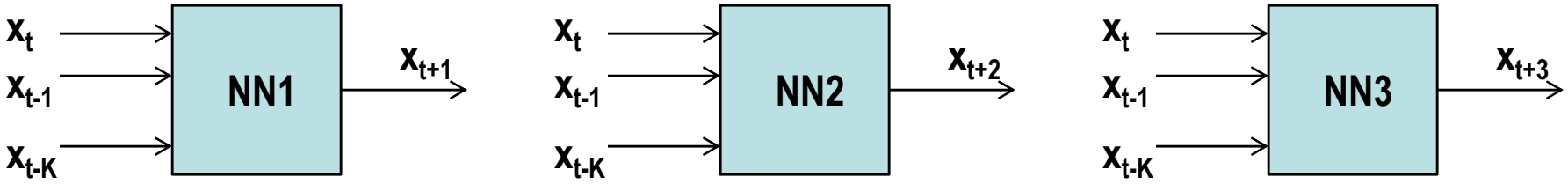
- We forecast several steps ahead, e.g. x_{t+1} , …, x_{t+k}
- Three major approaches:
  - **The Iterative approach:**

    Develop a model to forecast 1 step ahead.

    Forecast one step at a time, each time using previous forecasts as inputs.

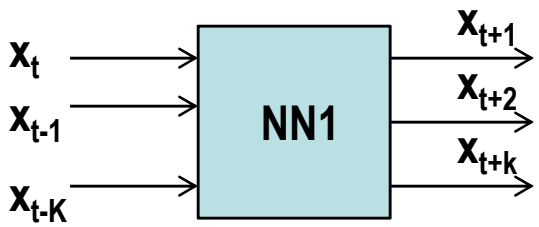
  - **The direct approach:**

    Develop a different model for each step ahead.



Multi-Step ahead Forecasting (Contd)

- Third approach:
  - **The parallel approach:**
    Develop one model that forecasts all steps ahead simultaneously.



- Best approach: parallel, then direct, then iterative.

Error Functions

- Mean Absolute Percentage Error (MAPE)

$$MAPE = \sum_{i=1}^{N} \frac{|\hat{y}_i - y_i|}{|y_i|}$$

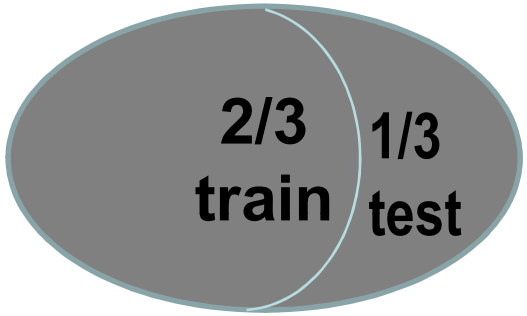
Where $\hat{y}_i$ and y_i are the forecasts and the true value, resp.

- Symmetric Mean Absolute Percentage Error (SMAPE)

$$SMAPE = \sum_{i=1}^{N} \frac{|\hat{y}_i - y_i|}{\frac{1}{2}(|y_i| + |\hat{y}_i|)}$$

Train/Test Partition

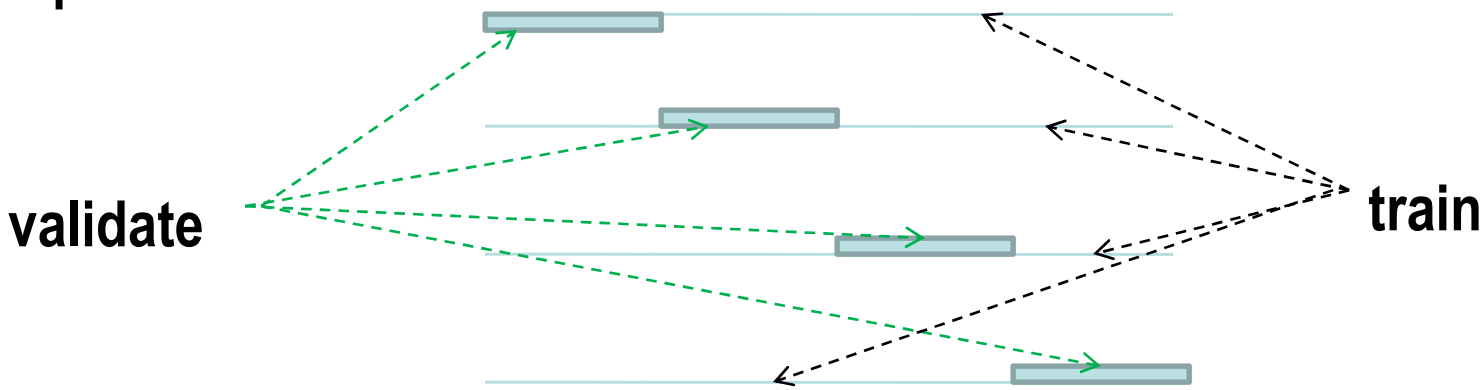
- Best practice:



- Test set should be used only once, at the very end of the design.

Parameter Tuning

- For every machine learning model there is a critical parameter that needs to be tuned, e.g. number of hidden nodes.
- Apply K-fold validation to the training set.
- Example K=4

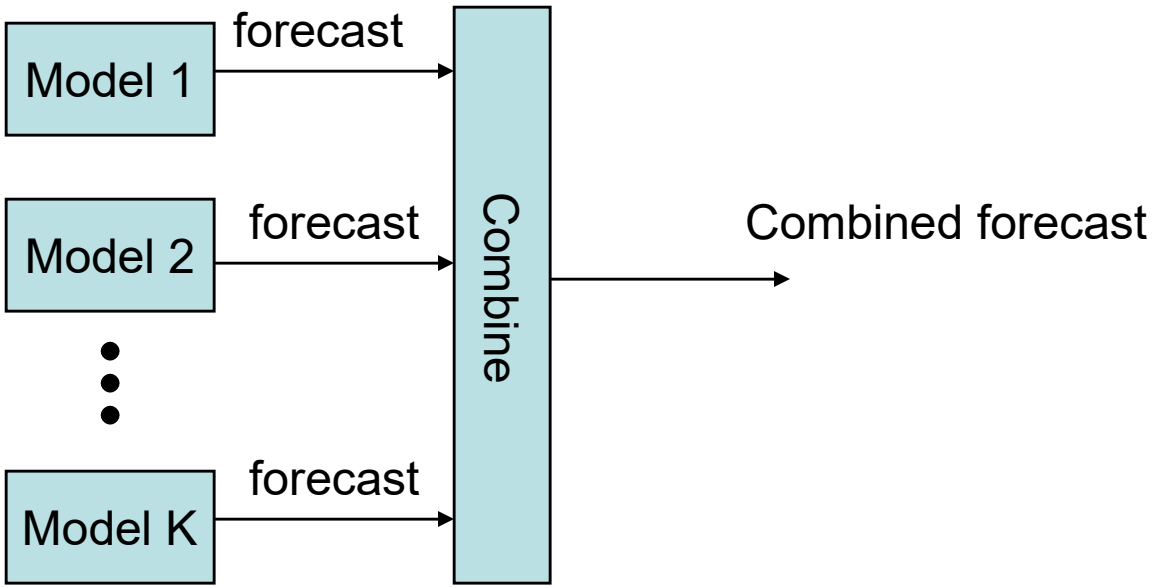


Overfitting

- It is a serious problem, and occurs if
  - We have a short time series.
  - We have a complex model with many parameters to tune.
  - We use too many lagged values as inputs.
- Try to use simple models.
- Beware of **forward looking**, i.e. tuning some parameters on the test set, or using the test set several times.

Forecast Combination

- It is recommended to develop several models, select the best few, then combine their forecasts.
- Do not combine all. Just the best few.
- The combined models should be as **diverse** as possible



Forecast Combination (Contd)

Final Forecast $y = \Sigma_i w_i y_i$ where w_i =combination weights

Methods:

- Equal weight: w_i = const
- Performance based (MSE_i is mean square error of model i)

$$w_i = \frac{1/MSE_i}{\sum_{j=1}^{M} 1/MSE_j}$$

- Rank based:

$$w_i = \frac{1/Rank_i}{\sum_{j=1}^{M} 1/Rank_j}$$

- Least squares: Obtain w_i by minimizing total error in y, using least squares.
- Nonlinear, e.g using geometric mean of forecasts.

M-Series of Forecasting Competitions

- M competition (1982), M2 competition (1993), M3 competition (2000), and most recently M4 competition (2018).
- These competitions offer thousands of time series, and teams participate to develop forecasting models.
- The teams are ranked, and we learn lessons on which methods or practices work best.
- Thorough analysis appears at the International Journal of Forecasting.

Lessons Learned

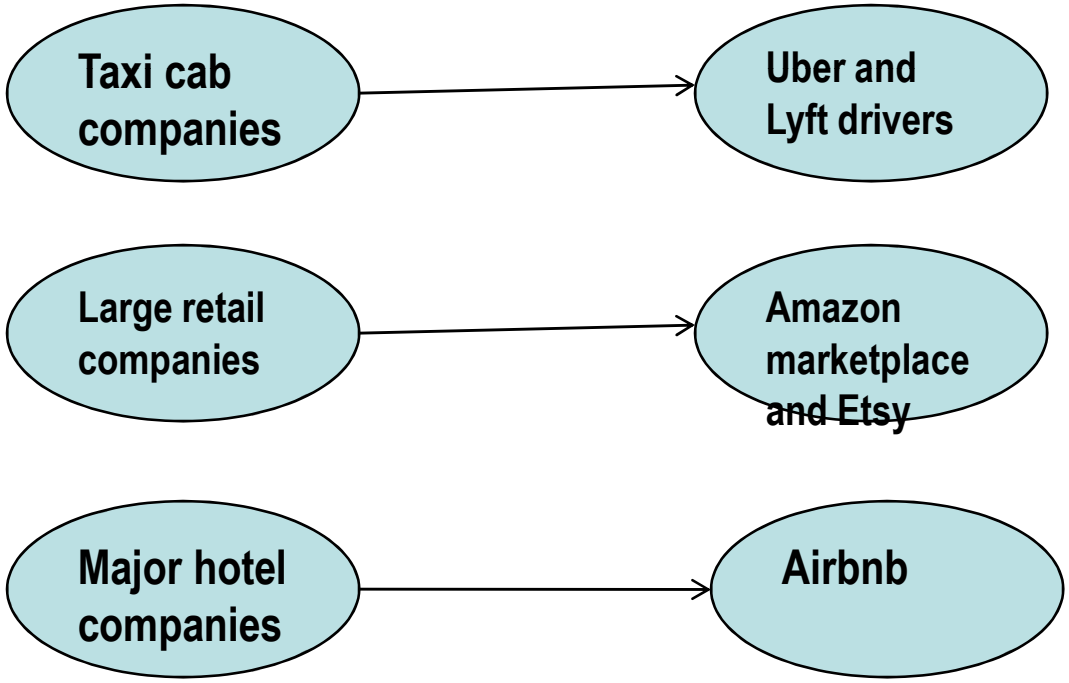
- Simple models typically perform very well, for example AR (auto-regressive), and exponential smoothing, especially for small size of data.

- Also, hybrid machine learning/simple statistical models perform very well.

- Forecast combination usually improves performance.

Opinion Mining for Hotels

- There is a trend in the last decade, that moves economic activity back to the small individual entrepreneur.



Social Media

- What moves customers to buy a product (other than price and quality):
- Advertising (controlled mainly by google and facebook).
- Online user reviews.

Social Media

- Opinions on social media have the same power as word of mouth, and even more.
- It has a grass-roots effect of shaping opinion,

Social Media

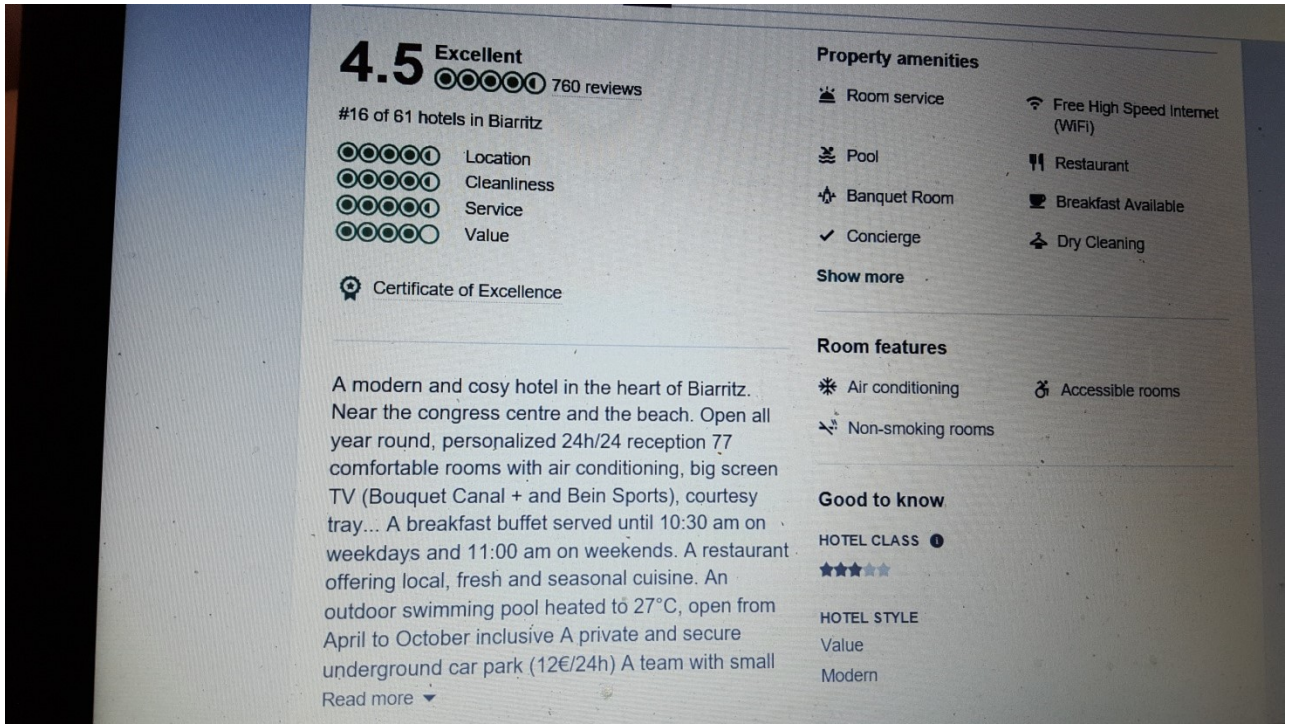
- The start:

- Now, many companies, such as

- And almost all vendors' sites.

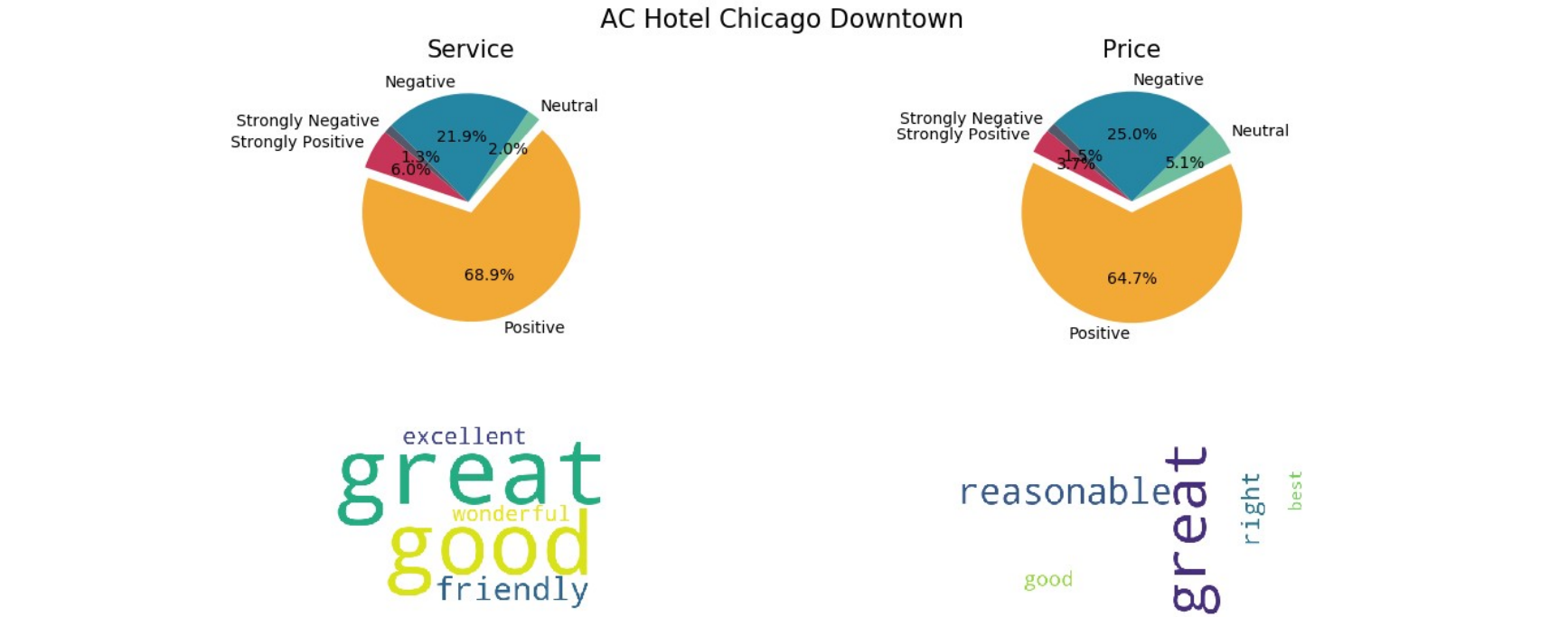
Hotel Sites

- Tripadvisor

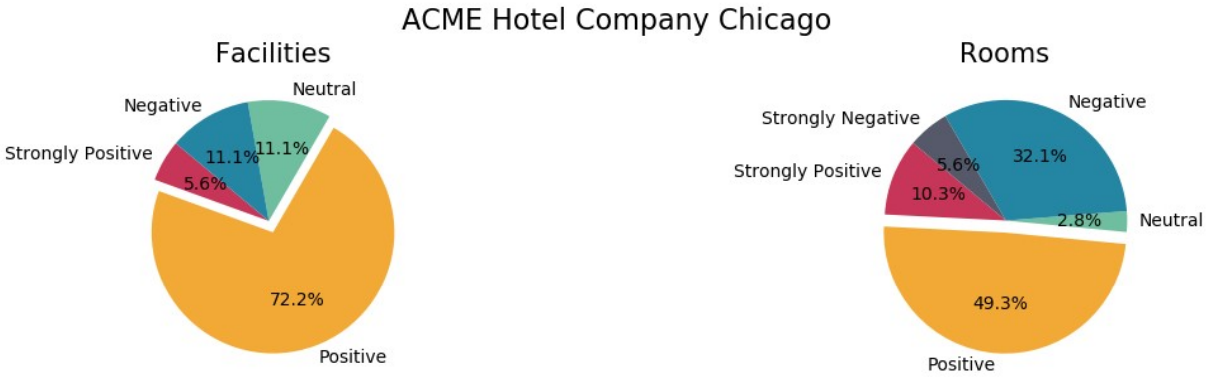


Hotel Sites

- More information can be extracted:



Hotel Sites


ACME Hotel Company Chicago
Facilities
Neutral
Negative
Strongly Positive
11.1%
11.1%
5.6%
72.2%
Positive
Rooms
Strongly Negative
Negative
Strongly Positive
5.6%
32.1%
10.3%
2.8%
Neutral
49.3%
Positive


close
great
pleasant
good
accessible


nice
comfortable
clean
great
small

Hotel Sites

Kristen greeted us at the front desk. She was so sweet and so nice at check in. Our room was clean. The rate was very reasonable. The bed was not as comfortable as I like. In the morning I spoke with the manager. He was very nice and apologized for the bed. I would stay here again. They also offer free coffee in the lobby in the morning.

'Our room was clean .'

'The rate was very reasonable .'

I stay in hotels every week as I work out of town. No fridge and no microwave and had to pay for wifi. I could have stayed at Comfort Suites with a fridge, microwave and wifi for the same price and will do so the next time. The A/C unit was messed up and loud. My company is in SC and gets a corporate discount and they wouldn't give on.

'I could have stayed at Comfort Suites with a fridge microwave and wifi for the same price and will do so the next time .'

Jessi at the front desk was incredible, very pleasant and professional. My flight was very late and she made my check in and stay free of stress and very easy. I definitely recommend this hotel and the restaurant was very good. Nina at the bar was very friendly and gave me a list of places to go check out in the area

'I definitely recommend this hotel and the restaurant was very good .'

The Advantage of Opinion Mining

- Too many reviews to read. It automates the process.
- It summarizes important points in the reviews.
- One can monitor the trend of opinions over time:
  - Average opinion was 8.2 last year, now it is 7.4.

Opinion Mining Approaches

- Lexicon based approach
  - Lexicon of good words, lexicon of bad words.

- Machine learning: input n-grams or word tuples.

Questions?

Please feel free to contact me any time during the year and further at amir@alumni.caltech.edu