

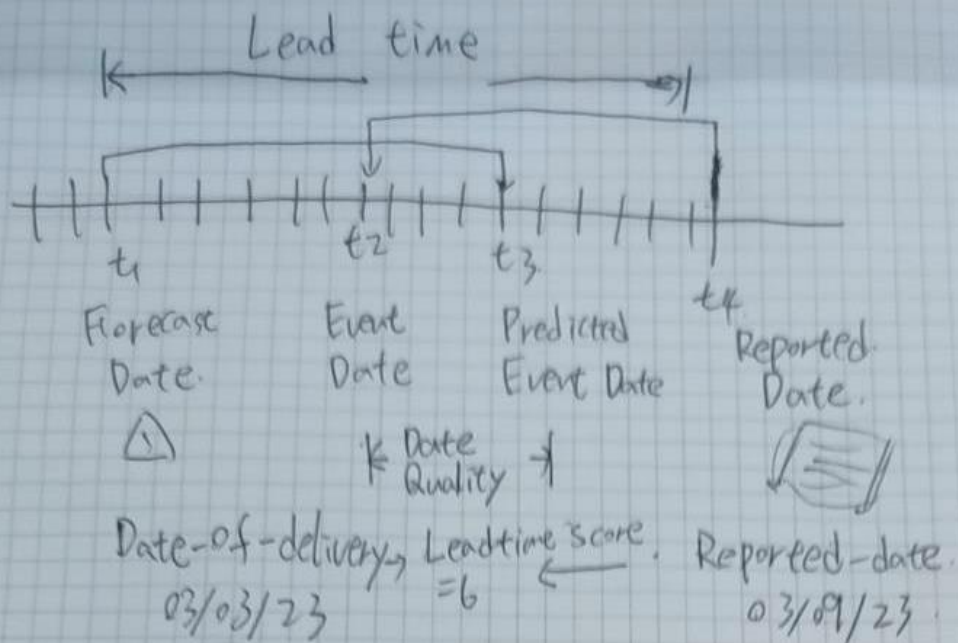
日期 = 6/3

題目 = Future Insights = Harnessing AI and Social Media
講者 = 美國維吉尼亞理工大學 呂昌田教授 for Advanced Event
and Epidemic Forecasting

- Forecasts automatically emailed for evaluation without human-in-the-loop.
- Evaluation done independently of authors, by MITRE
 - Using a Gold Standard Report (GSR) of events (from influential newspapers outlets).

Metrics

1. Quality (Graded on 0-4 scale).
2. Lead time (how far in advance?).
3. Recall
4. Precision
5. Probability (reliability).



Content Filtering via Dynamic Query Query Expansion

• Traditional Solutions:

1. Keyword match based on a predefined dictionary
2. Train a classifier (e.g., SVM) [www-2016].
3. Rules or decision tree based.

• Challenges for evaluating Twitter data.

- Heavily informal and ungrammatical.
- Huge Volume (more than 1 million per day).
- Dynamic (e.g., new hashtags, new topics).
- Noisy and unbalanced (e.g., less than 5% related to civil unrest).
- Rich social network structure.