# Exhibit\_F\_A

1. **Automatically detect spikes in existing discussions along with new topics for discussion.**

## Our interpretation: Detecting spikes refer to finding anomalies—First it is important to define an anomaly. Refer to example section for different kinds of anomalies.

Example: Anomalies can be broadly categorized as:

1. **Point anomalies:** A single instance of data is anomalous if it's too far off from the rest. *Business use case:* Detecting credit card fraud based on "amount spent."
2. **Contextual anomalies:** The abnormality is context specific. This type of anomaly is common in time-series data. *Business use case:* Spending $100 on food every day during the holiday season is normal, but may be odd otherwise.
3. **Collective anomalies:** A set of data instances collectively helps in detecting anomalies. *Business use case:*Someone is trying to copy data form a remote machine to a local host unexpectedly, an anomaly that would be flagged as a potential cyber attack.

## How will we achieve this?

* 1. First determine what common topics of conversation mentioning CMT are. This is a clustering task, since we do not have prior knowledge of what exactly the different clusters may be—SO, we want to implement some kind of unsupervised learning algorithm. This can help us learn what the different clusters are.
  2. However, if we already know the topics, example: Say, there is an upcoming event and most discussions surround the event, then we can move on to anomaly detection w/o any clustering.
  3. Once topics of discussion have been determined, we can use rnns or lstms for time series anomaly discussion if we want to feed live data
  4. Another approach is to use knn or local outlier factor (LOF)

## Resources/Tools/References?

A: Papers

1. Long Short Term Memory Networks for Anomaly Detection in Time Series
2. Detection of Anomalous Drops with Limited Features and Sparse Examples in Noisy Highly Periodic Data

# Anomaly detection in online social network: A survey

### [Semantic Expansion of Tweet Contents for Enhanced Event Detection in Twitter](https://ieeexplore.ieee.org/document/6425790/)

B: Blog posts

1. <https://www.datascience.com/blog/python-anomaly-detection>
2. <https://blog.statsbot.co/time-series-anomaly-detection-algorithms-1cef5519aef2>

C: Others

### [**Social Media Anomaly Detection - University of Southern California**](http://www-bcf.usc.edu/~liu32/SMAD_slides.pdf)