News Event Prediction using Causality Approach

TEO WEN LONG

A 1-page abstract is a *movie (thesis) trailer*. Avoid summarizing your Introduction chapter. Focus on the problem statement, hypothesis/objective, research approach, quantitative validation summary, and implication of your findings. For Ph.D., emphasize on original contributions.

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time.

Acknowledgement

The Malay abstract is written as the sentence structure of the English abstract. All specific terms must be checked with Dewan Bahasa and Pustaka (<http://prpm.dbp.gov.my/>).

# Introduction

## Overview

Newspaper is an important part of our life as it is a printing media in which all information of either national or international news are published and delivered to the public every day. Henry Ward Beecher (1887), an American social reformer and well-known speaker once said, “Newspaper is a greater treasure to people than uncounted millions of gold.” Articles within the newspaper play essential role in education development and makes public aware about events happening in the region or nation that they are living in. By reading newspaper, readers can learn and observe others point of view as it brings another whole new perspective on same events. In the digital era, artificial intelligence (AI) expert make use of the digital technologies to get a better insight within the traditional media, newspaper. Within these digital technologies, event prediction held a big portion as it forecast future events and it is valuable to alert public on predicted events.

Event prediction is a data analytic technique that make use of experience and knowledge as well as pattern from past to predict future events. Natural disaster prediction is one of the examples in applying concept of event prediction. For example, Japan is a well-known earthquake active country as its archipelago is in an area where multiple continental and oceanic plates collapsed together. Hence, earthquake forecasting is important for Japan and scientific report stated that earthquakes cluster in time and location as it can be predicted and take precaution before tragedy happened. Besides, event prediction is also useful in business intelligence. A report in 2016 also showed that business analysis and prediction help them to understand more about their customer, in order to enhance the success of their marketing strategies.

Every country had its obligation to protect its national security. National security is a requirement to maintain the survival of a country though economic, diplomacy, political and ethical power and focus on freedom from military threat and political coercion. Without national security, a country might at risk and attacks such as terrorism, sabotage, information warfare, etc might infiltrate the country. For example, ISIS threat stunned the world as a gunman, Mehdi Nemmouche opened fire at Jewish Museum of Belgium in Brussels as he is suspected in joining extremist groups, ISIS in Syria. This event took 4 lives of innocents.

South China Sea (SCS) is a conflict zone whereby an estimated USD5 trillion worth of raw products shipped through shipping lanes in SCS each year and its nearby countries made them fight over each other to have the main control of the whole SCS. The conflict is known as South China Sea disputes. The events of territorial disputes of South China Sea populated all the newspapers and many events regarding to the disputes were reported through national news agency. It rises concerns about the beginning of world war as for example a near-collision between US warship, Decatur and Chinese Luoyang missile destroyer in South China Sea highlights the escalating danger of confrontation between US and China. (Ni, 2018) SCS is strategically located at peripheral ocean that is a piece of the Pacific Ocean, starting from Karimata and Malacca Straits to the Strait of Taiwan with area about 3500000.Besides, South China Sea is rich in marine life and natural resources such as oil and natural gas, even have the most of the important shipping lanes in the world.

Causality is the relationship between cause and effects. Every event will occur first on cause and followed by effects. In SCS disputes, causality is highlighted between benefits from SCS (cause) and territorial disputes (cause) is clearly highlighted in SCS disputes events.

In order to have an advanced insight among these disputes, event prediction is necessary, and causality should be taken as main attributes. However, there are still several problems and challenge to be solve in order to achieve an excellent prediction model based on causality.

## Problem Background

National security is always the top priority of governments to protect society from disruption owing to a disaster or crisis. There are many aspects on national security such as territorial, economic, physical, social, political etc. However, the peaceful of national security had been affronted by South China Sea disputes.

Due to geological and resources advantages of South China Sea, countries within the region such as Brunei, China, Taiwan, Malaysia, Indonesia, Philippines, Vietnam etc. made competing the territorial claims over it. Based on news on The National Interest in 2016, an estimated US 5 trillion worth of global trade passes through the South China Sea annually. Hence, territorial disputes in the South China Sea started to concern worldwide community about peace of world. In order to claim the ownership of South China Sea, countries are challenging against each other by putting military force in the area. This can be observed from the news of China spent almost 1 year to build 7 new islands by moving sediment from the seafloor to reefs and after that focused on building ports, airstrips and other military structures on the islands.

South China Sea dispute had a brief background involving timeline from 221 BC until recent. Each of the historical event occurs and accumulates and eventually things go haywire. Many dispute events happened in either small or large scale. For example, Spratly island dispute and “nine-dash” line that proposed by China is some of the significant disputes in South China Sea. Besides than these two issues, there are many issues that remain unsolved and will constantly concerning the worldwide community.

However, all the information retrieved from news article are unstructured. Unstructured data have no recognizable structure via pre-defined data models and schema and mainly generated by human or machine. . By collecting these unstructured data from the past and analysing its trends, we are able to have a better understanding about what may happen in the future. . In SCS disputes, event prediction is important to give public a better understanding about future events that might happen. A better policy can be made with regards to protect national security under SCS disputes with the event prediction technique based on unstructured data in news articles.

In event prediction based on news articles, there may have some challenges. First, news article is a type of unstructured data that contain a lot of valuable information in term of cultural, social and historical but doesn’t fit into traditional row and column structure of relational databases.. It require substantial manual effort to analyse and extract the essential information from news articles. Second, a event that causes another events may completely different from the real prediction. It is indicating that the predictive model provides a faulty outcome that hard to distinguish from the true prediction.

There are several researchers research on topic event prediction with different method. Granroth-Wilding (2016) proposed a predictive neural network model that learns embeddings for words describing events, a function to change embeddings into event representation and a function to predict the degree of relationship between two events.However, the model is more focus on chain or events sequence which is good for rich-infomrative events but might not suitable for news articles that have unordered sequence. Preethi (2015) proposed an event prediction model for Tweets using temporal sentiment analysis and causal rules extraction. This model is useful to analyse user’s sentiments and predict future events using temporal attribute. This study analyse sentiments of user’s opinion and is not suitable for news articles whereby formal news reports seldom express their sentiments within the articles.

Radinsky (2014) proposed a predictive model based on causality attribute. The predictive model is suitable for unstructured data such as news article and learn from previous events to predict the future. For example, Given a South China Sea events stated that “China (Actors) construct (Actions) a new military purpose platform (Objects) on strategically-located Bombay Reef, South China Sea (Location) in 22 Nov 2018 (Time)”. From this event, we can observe clearly that there are 5 tuples <Actors, Actions, Objects, Location, Time>. From the 5 tuples, we can extract useful keywords that identify the characteristics of the events. The further event happened after this is that “Vietnam’s sovereignty in South China Sea is threatened and Vietnam started to protest”. This causality pair indicated the relationship between cause “Country expand their territorial area in South China Sea” and effect “Another country’s sovereignty is threatened and started to protest”. This model makes use of previous events to predict future events based on causality attributes.

## Problem Statement

South China Sea (SCS) is a conflict zone where events happened from time to time with different severity. This greatly impact or influence the policy that made by government of Malaysia to overcome the negative impacts brought by SCS territorial disputes in term of national security. The problem is to extract valuable information from SCS conflict events and predict the future events that may happen. Besides, online news is unstructured data and extracting correct information from massive online news articles that contain different resources automatically is part of the challenges. Event prediction is biased to measure, and a good predictor is needed to make sure that its prediction is accurate and precise.

## Aim/Purpose of Study

This study will address the SCS disputes issues by implementing Radinsky (2012) predictive model using Pundit algorithm. This predictive model learn causality of events and provide accurate suggestion based on predicted outcome.

## Objective

1. To extract keywords from news articles using 5 tuples approach on <Actors, Actions, Objects, Location, Time> for better event representation.
2. To produce generalized node and path from extracted keyword and finding minimal generalization path for generalizing extracted information.
3. To produce abstraction tree (AT) based on events nodes and causality prediction rules in order to building a suitable prediction model and evaluate the accuracy of prediction model with precision and recall or comparing with human predictor.

## Scope

Event prediction is important to have better understanding for future event that may happen, but it also have its limitations that need to address and narrow down. Below are the scope of this study:-

1. The coverage of this study will be using news articles that extracted related to SCS conflict from the national news agency.
2. The study will focus on SCS conflict news from Vietnam News Agency (VNA) and Xinhua News Agency (China)
3. Prediction model in this study will be based on Pundit Algorithm proposed by Radinsky (2012).
4. The study will only covers the causality prediction rules that adopted from Radinsky (2012).
5. 5 tuples approach will be used to extract the keyword on <Actors, Actions, Objects, Location, Time>

## Significant of Study

Event prediction is important because it give a better insight and forecast for public about events that possible to happen. Besides, in SCS conflict, event prediction can be used to protect national security of Malaysia and appropriate actions can be took by governments to prevent the happening of unseen tragedy based on predicted events. In addition, Malaysia governments are able to protect the sovereignty of country in SCS region to claim peace and avoid involving in SCS conflict.

## Organisation of Study

**Chapter 2** will be discussed about the literature review of event prediction and information extraction from news articles based on different method and referenced method. **Chapter 3** will focus on research methodology, research framework, overall research phase, and measurement and rules. **Chapter 4** will be present about initial result of implementing prediction model proposed by Radinsky (2012) by on step-by-step approach on causality attributes.

# Literature Review

## Introduction

This chapter will discuss the literature review of event prediction using causality approach on South China Sea (SCS) conflict. First, the chapter will briefly explain about event prediction and SCS conflict Then,the chapter will discuss text mining which consists of many stages such as text preprocessing, event extraction, event representation, event generalisation and event relation. After that, this chapter will briefly discuss the prediction algorithm and justify the most suitable algorithm for SCS conflict event prediction in this paper. Finally, this chapter will compare existing work that had done in the domain of event prediction. Last but not least, this chapter will discuss open issues in event prediction that need to be addressed, followed by a short brief summary.

## Event Prediction

Research on areas of event prediction are wide as it involves different domains such as natural disaster of earthquake and tsunami , political and sociality , economics , and etc. In this paper, event prediction is focused on textual events from news articles. Figure 2.1 shows taxonomy of event prediction. In this taxonomy, the flow of event prediction is clearly illustrated and started with general text mining, followed by choosing clustering algorithm, determine event relation with either temporal or causality relation and apply prediction algorithm to build a predictive model as well as train with extracted data. Eventually, prediction validation will be perform to measure the performance of the model based on accuracy and reliability.

Taxonomy of Event Prediction

Taxonomy of Event Prediction

## South China Sea Conflict

South China Sea (SCS) is located strategically within Asia country such as China, Taiwan, Indonesia, Philippines, Vietnam, and our country Malaysia. It has the most busy shipping lane as one-third of world’s shipping passes through SCS and almost 3.37 trillion global trade happened within SCS in year 2016 . Due to these huge market profit, many country started to claim that SCS is their own territory and China even illustrated a "nine-dash line" which is a huge part from SCS and claim that region within "nine-dash line" is China’s territory . These started the dispute between countries. For every disputes, public are able to know the flow of events through newspaper. National news agency such as Xinhua News Agency for China and Vietnam New Agency (VNA) stands the voice of government. stated that Xinhua is the biggest propaganda machine that spread the core concept of China’s government to the public. Hence, there are huge valuable information that able to obtain within news article among these national news agency.

## Text Mining

Text mining, also known as knowledge discovery from textual databases and related with steps of obtaining valuable and non-trivial pattern or knowledge from unstructured text documents. . A general framework proposed by shows that text mining contain two main elements, which are text refining and knowledge distillation. Text refining are techniques to transforms free-form text into intermediate form meanwhile knowledge distillation is to obtain the pattern or knowledge within the intermediate form. It is illustrated in Figure 2.3.

General Text Mining Framework

General Text Mining Framework

However, there are open problems in text mining that need be addressed. For examples, variety of intermediate form that causes complexity and uncertainty in text mining, multilingual text refining, domain knowledge integration , and customized autonomous mining .

### Types of Data

Nowadays, volume of data is dramatically increasing, associated with high flow of data and high variety of information. It is known as big data. Big data consist of 2 type, structured data and unstructured data.

#### Structured Data

Structured data is refer to the data that organisation all the information in formatted way and easily to extract and analysis by relational databases. . Example of structured data are phone number, name, identification number, date,etc. Format of structured data is analysable with human-generated queries and algorithm.

#### Unstructured Data

Unstructured data such as text, emails, blogs, web pages, comments on social media is getting enormous nowadays. Figure 2.1 shows several types of unstructured data that common in our daily lives. Information contained within these text originating from humans is absolutely richer and more valuable than traditional numeric data sources . Besides, study in 2012 showed that about 80% of enterprise data are unstructured. Because of that, exploring and extracting useful information from these unstructured data, also known as text mining is essential. However, text mining is difficult due to ambiguity of languages, multiple duplicated words with different meanings, abbreviation, linked references etc. In Figure 2.2, general type of unstructured data is showed graphically for better understanding.

Type of Unstructured Data

Type of Unstructured Data

To overcome the difficulties, many method such as Named entity recognition (NER), Nature Language Processing (NLP), sentiment analysis, causality relationship extraction etc. are developed to extract high quality information from textual data.

### Text Preprocessing

#### Tokenisation

#### Stop-words Removal

#### Lexicon Normalization

taxonomy - POS tagging, stemming, lemmatisation

#### Chunking

### Event Extraction

As increasing amount of digital data available nowadays, text mining is essential to extract valuable information from unstructured data. Event extraction is one of the common application of text mining as it obtaining specific knowledge concerning incidents referred to texts. . Event extraction had been involved in different tools and application. For instance, information from electronic medical records can be extracted automatically based on professional clinical advices to improve medical cases detection. . Besides, had proposed an automatic crime prediction from twitter posts using event extraction. It is beneficent to defend against happening of crime incidents.

Based on report by ,there are three main approaches to event extraction, which are stated as below:

#### Data-Driven Event Extraction

Data-driven approach is mainly used for natural language processing (NLP) application and focus on quantitative methods. One of the weakness of data-driven approach is that it doesn’t consider about meaning and sentiments and require a large amounts of data to get statistically significant results.

#### Knowledge-Driven Event Extraction

Compared to data-driven, knowledge-driven is more focus on patterns and exploring human-knowledge. It considers sentiments and consumes less training data compared to data-driven, and also able to define powerful expression by using semantic elements for an easily interpretable and traceable result. However, knowledge-driven requires more advanced knowledge and more difficult to extract knowledge from unstructured data.

#### Hybrid Event Extraction

In order to balance both advantages and disadvantages between data-driven and knowledge-driven extraction, hybrid approach is used to overcome the lack of expert knowledge in pattern-based problem, by applying statistical methods. One of the weakness of hybrid approach is that it required expertise in dealing with complex extraction techniques.

### Event Representation

After extracting useful information from event extraction techniques, these data is match with certain meaning to represent certain knowledge. In work of , they first use ReVerb ,a autonomous program that identifies and extract binary relationship from English sentences with tuples of event, then use ZPar to parse the sentence with subject, object and predicate. The event is represented as tuple *E = (O1,P,O2,T)*, where O1 is the actor and O2 is object, P is the action and T is the timestamps of the event.

Event representation using tuples is widely used with different domains in different paper . By using tuples, it provides clearer picture of the event and easier to categorise each element in tuple with their corresponding category.

#### Sentence Level Description

#### Syntax-driven

#### Semantic

5 tuple

### Event Generalisation

Generalisation on causality pattern is very useful in discovering high-level causality rules behind specific causality pairs. . Given an event "*a massive 8.9-magnitude earthquake hit north-east Japan on Fridaya large amount of houses collapsed*" can be generalised into *earthquake hithouse collapse*. It is useful to obtain a general yet simple causality patterns.

Besides, utilised the concept of generalisation by finding the similarity between events. HAC hierarchical clustering is used to cluster all the generalised event pairs to build a hierarchical graph of events, called abstraction tree (AT). Without generalisation, the branch of AT will be huge and impossible to derive prediction from such as huge AT.

#### Generalisation Path

minimal generalisation path - measure semantic relatedness

#### Clustering Algorithm

Clustering is a process to divide huge volume of data into number of groups in such that each group have the most similarity than those in the other group. . Clustering is important for identifying the characteristics of unstructured data and tag them with corresponding category. There are commonly four type of clustering, which are hierarchical, distance-based partitioned and organizing map algorithm.

1. **Hierarchical Clustering**

* Hierarchical clustering uses the concept of successively merging each data into predefined cluster based on their similarity. It is represented by nodes and connectors and a sample of hierarchical cluster is illustrated in Figure 2.4. There is 2 types of hierarchical clustering, agglomerative and divisive. Agglomerative is a bottom-up approach which the algorithm starts with a single objects as separate category and repeatedly merge with another similar category until meeting stopping condition. Divisive is a top-down approach in which the algorithm start with one cluster and splitting recursively into branches. The advantages of hierarchical clustering is simplicity, flexibility and use similarity as measurements while its disadvantages is that stopping condition is hard to obtain. In the study, hierarchical clustering will be used because abstraction tree (AT) is produced with adoption of causality graph.
* Hierarchical cluster sample
* Hierarchical cluster sample

1. **Distance-based Partitioned Clustering**

* This type of clustering make use of distance measurements to classify homogeneous data. Depart from hierarchical clustering, the problem of local minima that caused by branches of hierarchical cluster can be eliminated. K-mean clustering is one of the example that used Euclidean distance to calculate nearest mean between two cluster. Figure 2.5 illustrated sample of K-mean clustering algorithm that differentiate different cluster according to their similarity and plot onto the graph.
* K-mean clustering
* K-mean clustering

1. **Organizing Map Algorithm**

* Self-organising maps (SOM) algorithm is one of the organizing map algorithm that produce low-dimensional regular grid and can be effectively utilized to visualize and observe characteristics of data. . SOM use concept of neural network with neuron and neuron weights are adjusted based on their features towards "winning" neurons. Figure 2.6 shows a visualisation of SOM. SOM comes with two major advantages, which are reducing data dimensions as SOM requires no target vector and without any external supervision and easily to classify based on data similarity.
* Self-Organising Maps
* Self-Organising Maps

### Event Relation

Event relation is important in text mining as it provides better insight and understanding about event happens either before or after. Generally, there are 2 major type of event relation that can be observe and obtain from news article which are temporal relation and causality relation. As events happen from time-to-time, temporal attribute should be focus as it measures the trend and evolution of event. Besides, there is always a cause for events to happen, hence causality is used to relate events from the very beginning based on its cause.

#### Temporal Relation

Temporal is indicated as time-based or using time as the main measurement. Usually, events getting worse from a small case. For example, in medical domain, long-term follow-up is important for medical investigation. make use of temporal relation to observe the trend of patient’s health condition over time, and eventually predict patients’ risks for a future adverse outcome. Besides, temporal relation had been widely adopted in event prediction with different domain such as micro-blog and tweets , sports , stock market flow and etc.

#### Causality Relation

The main concept of causality is that one or more things/events as causes can cause one or more things/events to happen as effect. In news article, there are many events that related with each other but hard to be observe and detect by human or massive news events happen in the same time cause human hard to digest all the news happened in one time. Thus, people tend to automated the extraction of news article based on causality. However, there are many challenges in automating the process. For example, faced difficulties on explanation knowledge graph through causality extraction from text such as causal-boundary determination and effect-event pattern determination. Nevertheless, it is useful for event extraction because it contain cause and effect that allows people to understand more about the following events. In this study, we will focus on causality relation as SCS conflict happens after certain small case instead of following time flow.

1. **Causality Pair**

* Causality pair is also known as entity pair, used to represent causality within pairs. For instance, stated possible causality pairs in his study as (1) main events of consecutive sentences, (2) pairs of events in the same sentence, (3) an event and a time expression in the same sentence. If the pairs are tagged as (*e1,e2*), the pairs will be *events-events*, *events-timex*, *timex-timex*.

1. **Causality Rules**

* After identify the potential cause-effect pairs, it is difficult to execute human-annotation on these causality pairs as the amount of data is getting enormous. Hence, autonomous annotation is proposed. . In order to get an accurate categorised result based on causality pair, causality rules is set in every text extraction. construct a set of causality rules that follows the template of <**Pattern,Constraint,Priority**> where Pattern indicates the trend of events between causality pairs, Constraint indicates the limitation on sentence which the pattern can be applied, Priority indicates the priority of causality pairs to be execute first or second.
* In the work of , Radinsky use causality rule to predict the future event as abstraction tree (AT) had been build earlier and using causality rules, a causality graph is built to explain the predicted event clearer.

1. **Causality Graph**

* Causality graph is a path diagram that visualise the causal relationship between events and enable researchers to estimate effects sizes and costs . Generally, Granger causality is widely used in causality graph for event prediction model as it is a statistical concept of causality. . In Granger causality, a true cause-and-effect relationship is obtained from particular variable that comes before another and it used bottom-up approach to predict the outcome. In , causality graph is used to illustrated relationship between 2 events. From Figure 2.8, an event "US army bombs a weapon warehouse in Kabul with missiles" is used and the event is extracted based on cause and effect and details are stated in causality graph.
* Example of Causality graph, adopted from
* Example of Causality graph, adopted from

## Prediction Algorithm

Prediction algorithm is an algorithm that modelling prediction and provides a guideline throughout the process of building up a predictive model. After obtaining extract keyword from textual data and clustering them based on clustering algorithm and relation extraction, the data is now said to be clean and sanitised. These clean data is then involve in prediction algorithm to build up a prediction model.

### Sequential Prediction Algorithm

Sequential prediction is a fast and simple pattern matching algorithm whereby comparing the past data or experience on a linear running timeline. There are many examples for sequential prediction algorithm such as medical condition that occur over time and patient conditions is getting either better or worse based on previous event. Besides, suggestion system that used in music application such as Spotify also involve in sequential prediction.

### Neural Network Prediction Algorithm

The biggest advantages of applying neural network in event prediction is the high tolerance and acceptance ability of noisy data and high accuracy. Despite of its complexity and long training time, many researches are more towards to neural network.Neural network use the concept of brain metaphor for information processing and is suitable to discover previously unknown, valid patterns and relationship in large data sets.

### Decision Tree (Pundit) Algorithm

Pundit algorithm is used in to learn and predict causality. It is one part of decision tree algorithm. Figure 2.8 shows the structure of Pundit algorithm and it started from extracting historial news article into causality pairs. Then, it generalise the pairs using causality prediction rules and create an abstraction tree for prediction. When a new event come in, the events will compare the bottom node with similarity and after rounds of comparing, an effect event is produced and take as predicted event. Due to the highly flexible and fix with causality from news article, this study will focus on using Pundit algorithm to predict the future events.

Structure of Pundit Algorithm

Structure of Pundit Algorithm

## Existing work of Event Prediction

There are several researchers research on the topic of event prediction based on causality attributes and combining different method on different domain. Table 2.1 shows the existing work of event prediction, focusing on textual data from different domain such as financial, social media, medical etc.

Existing work of Event Prediction

|  |  |  |  |
| --- | --- | --- | --- |
| **Author, year** | **Work** | **Method** | **Limitation** |
|  | Sequential event prediction | Given a sequence past database to predict next event within a current event sequence | Rely on additional user input, supervised ranking algorithm which not suitable for less predictive power past events |
|  | Narrative event evolutionary graph (NEGG) for script event prediction | Extract narrative event chain from newspaper and construct NEEG based on the chain |  |
|  | Depp learning on event-driven stock market prediction | Extract news text with dense vectors and train with novel neural tensor network, provides prediction based on convolutional neural network | Outcome is more focus on statistical, less focus on financial news that affect the share price in the market |
|  | Temporal sentiment analysis and causal rules extraction from tweets for event prediction | Event prediction based on temporal sentiment and causality from tweets which are from the opinion of people | Sentiment from tweets is uncertain and may affect on the predictive output |
|  | Event Prediction using Compositional Neural Network Model | Use word2vec to represent events and build a compositional neural network model for future event prediction | Multiple event chain in CNN models and limited to learning associations between pairs of events |
|  | Crime event prediction with dynamic features and matrix factorisation | Dynamic features such as visitor entropy, visitor homogeneity, region popularity, visitor ratio and count, observation frequency is taken as measurement in crime event prediction | Different city have different crime data, result may be biased |
|  | News events prediction using Markov logic networks | Markov logic network is used represent complex events by first-order logic and binded with domain-specific causal rules. Web ontology language (OWL) is used to perform causal inference and augmented by Markov logic network which uses weighed first-order formulas to represent probabilistic knowledge |  |

## Open Issues and Challenges

Event prediction is important and useful for public to have better insight and quick response against event might happen. However, there is several issues and challenges need to be addressed in order to produce a perfect predictive model approaching human-thinking. The issues and challenges are:

1. **Low performance of predictive model due to low resource scenarios**

* For predictive model, it require as much data as possible to train and test its accuracy and improve any weakness that learning from bad data sources. However, one of the problem is that there are scenarios of low resource that the prediction model can obtain. For example, if the model receive data source that are unique and only few events occurs because of these, the predicted output might not be valid and match with human-predicted output.

1. **Word Sense Disambiguation**

* In textual data extraction, word sense disambiguation is an open problem that troubling in determine the sense of word used in a sentence. One of the significant example can be observe in differences between dictionaries. For example, the word "**bank**" has 3 different meaning according to Cambridge Dictionary. First, bank is indicated as financial organisation where people can invest or store money inside, second is sloping raised land along sides of river, third is related to row of similar things. There are 2 sentences, "The **bank** will not be open in Saturdays" and "the river overflowed the **bank**.". The ambiguity between 2 same words but different meaning make machine difficult to detect and extract correct information from its true nature.

1. **Pronoun Resolution**

* Pronoun resolution also known as anaphora resolution, is an common problem of resolving what a pronoun, or a noun phrase refers to. . Taking example of "John helped Mary", followed by "He was kind". As normal human, we can understand that "he" is indicating to John but not the machine. Machine will automatically define "he" as a new character instead of combining John with "he". Anaphora resolution is still an active research which contributes to better machine learning technique.

## Summary

In this literature review, concept of event prediction and South China Sea conflict is clearly explained before moving in brief stage of event prediction. Then, stages within event prediction such as text mining, clustering, event relation extraction, prediction algorithm and etc is mentioned with details in the previous subchapter. Finally, this chapter discuss about previous and existing work that done by other researchers on the textual data event prediction in different domain and also the open issues and challenges in event prediction so that reader get better understanding about constraints of event prediction nowadays.