

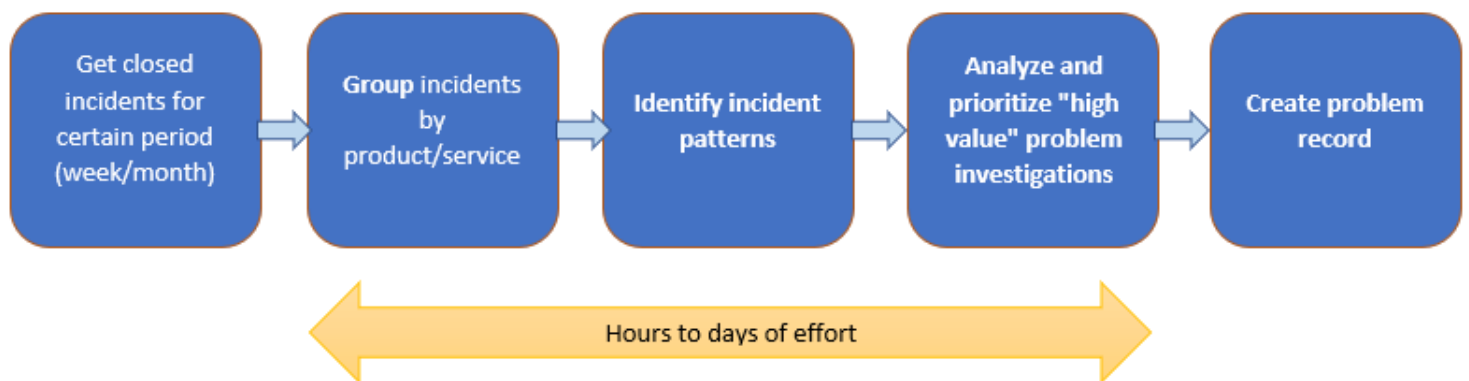
# IDENTIFY PROBLEMS FOR INCIDENT MANAGEMENT USING AI/ML



## Introduction

The primary goal of the problem management process is to prevent recurring incidents. However, identifying critical problems from thousands of incidents is a tedious and error-prone activity. Let's understand how artificial intelligence/machine learning (AI/ML) technologies can help us identify problems more easily.

Let's understand the problem identification process flow.



# Challenges faced by problem managers

Incident information that describes the problem faced by the end user is textual data, and identifying patterns in that data using existing reporting or analytical tools such as Excel is time-consuming and error-prone.

When identification of patterns is not complete (a highly likely outcome when humans are responsible for recognizing patterns in data), a lot of problems may go unnoticed.

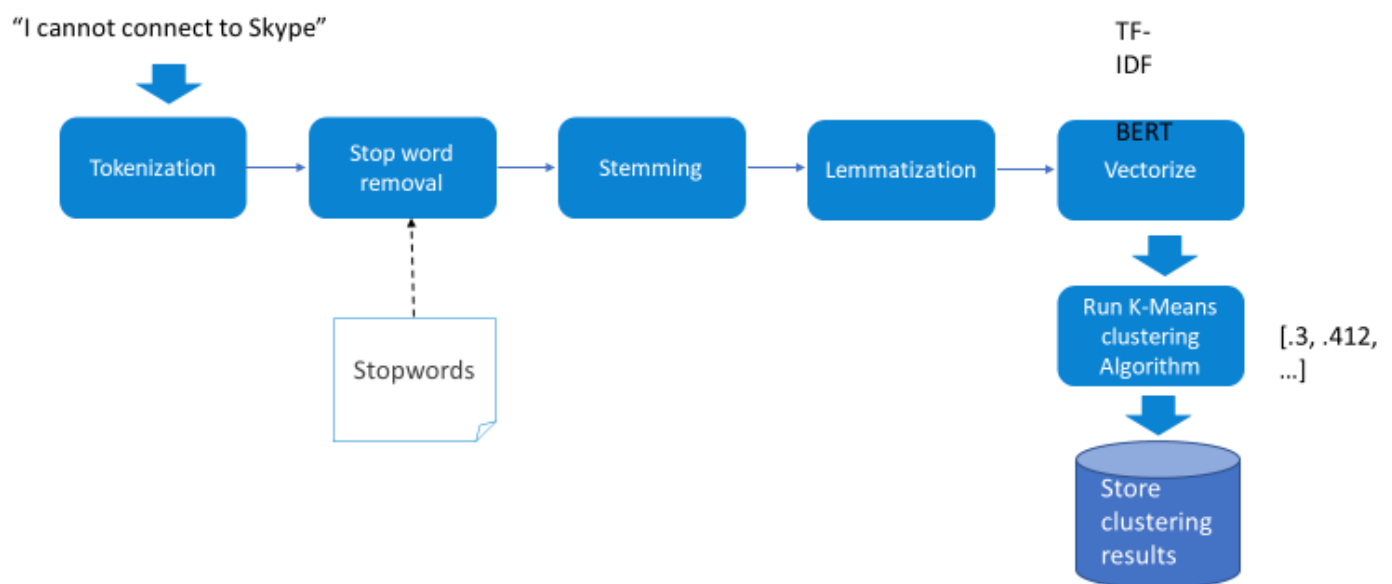
## How can AI/ML help solve this problem?

Because the incident data is textual in nature, we need a natural language processing (NLP) module to process it, which will remove stop words (common words/filler words), help identify the intent of the textual data, and convert it to numerical data.

With numerical data, clustering algorithms such as k-means clustering can be applied to group similar data points together.

Let's see how this data processing pipeline looks.

### Data Processing Pipeline



© Copyright 2021 BMC Software, Inc.

1

Let's understand each step in further detail.

## Tokenization

In this step, individual words in the sentence are separated. This also separates any punctuation and word separators such as dashes or semicolons.

## Stop words removal

In this step, filler words or common words are removed. For example, company names or words like “prod” or “dev” may be appearing in incident text repeatedly, making pattern identification for problems more difficult.

## Stemming

In this step, affixes to words are removed to get to the base form of the word. For example, words like “working” or “worked” will be reduced to “work.”

## Lemmatization

This step is similar to stemming, but here the root word of the given word is found, whereas stemming just removes affixes. etc. For example, the root word for “is,” “are,” and “was” is “am.”

## Vectorization

In its simplest term, vectorization is used to represent text data in numerical vectors. Word vectorization is an NLP method for mapping words or phrases to a corresponding vector of real numbers, which is used to find word predictions and word similarities.

## K-means clustering

K- means clustering is an unsupervised learning algorithm where K represents a number of clusters. K-means performs the division of objects so that similar objects are grouped together as clusters.

Once a group of incidents’ summary/description is tokenized, preprocessed, and converted into its vectors, the K-means clustering algorithm groups similar vectors together, forming clusters of incidents that have similar text or intent.

## Problem identification

With AI/ML, we have greatly automated the analysis and identification of a problem. What took hours and days can now take minutes with AI-based clustering for finding patterns in the incident data to generate clusters. Once clusters of similar incidents (by intent) are created, the problem manager can easily perform further analysis (e.g., clusters with a high number of incidents or clusters with high effort or high time to resolve) to identify critical problems from incident information and prioritize the most high-value problems to investigate.



# Conclusion

AI/ML technologies can help IT problem managers to identify impactful problems quickly, reduce manual efforts, and isolate recurring problems that impact business and operational efficiency.

At BMC, with the vision to transform any business into an [Autonomous Digital Enterprise](#) (ADE), we maximize the use of AI/ML to make processes easier, enhance actionable insights, and increase organizational agility.

Our AIOps solutions apply machine learning and predictive capabilities across IT operations and DevOps environments for real-time, enterprise-wide observability, data-driven insights, and automated remediation.