**Required paper structure:**

*Headings (Project title, students’ names and ID’s, supervisor’s name)*

*Abstract and keywords*

*1. Introduction*

*1.1.Organization of the paper*

*2. Background and Related Work*

*3. …[additional section(s) as needed]*

*4. …*

*5. Expected Results*

*6.1.Requirements (Use Case)*

*6.2.Design (GUI, UML diagrams)*

*6.3.Testing plan*

*References*

To develop an app that can store, analyze and suggest solutions for every Mcare report, providing more efficient day-to-day and long term work.

* All logs are examined manually by call center
* All work flows and troubleshooting are based on representative’s experience
* Existing, reoccurring and resolved errors are verified manually
* To analyze the Mcare Daily report - provide more work efficient environment.
* To timeline and analyze all previous site cases to identify reoccurring issue on port \ switch per customer.
* To synchronize and monitor any frequent alarms\ triggers to locate reoccurring issues with Mellanox products on all Mcare Sites
* To provide a day \ week statistics - threshold adaptation advise for each site.

Abstract:

*In our project we employ methods of machine learning from text retrieval to search and build a workflow to troubleshoot a problem with a network based on previous similar errors in company database and what had been done to solve them providing more efficient day-to-day and long term work environment.*

*In addition the program will analyses the errors on time and configurational scale and advice on threshold adaptation to prevent future network issues and false alarms.*

*Text pattern recognition is a difficult task due to the wide array of technical errors and the conversation recognition to identify what steps are helpful and what are not.*

**Keywords***:*

*~~In our project we employ methods from text retrieval to search for and localize all the occurrences of an object in a picture data base, given a query image of the object.~~*

*~~Visual search of an object is a difficult task, due to changes in viewpoint, illumination and partial occlusion. Thus, the object is represented by a set of viewpoint invariant region descriptors so that recognition can proceed successfully. Efficient retrieval is achieved by employing methods from statistical text retrieval, including inverted file systems, and text and document frequency weightings. This requires a visual analogy of a word which is provided here by vector quantizing the region descriptors. The final ranking also depends on the spatial layout of the regions. Due to the massive groundwork the retrieval is immediate like Google's search engine.~~*

**~~Keywords~~**~~: object recognition, viewpoint and scale invariance, text retrieval, SIFT.~~

1. **INTRODUCTION**

The aim of this project is to retrieve comments in the correspondences between support engineers and customers from the company's database and identify helpful and unhelpful advices using Natural Language Processing (NLP) algorithms on both the engineer and the customer… and to build a multi-stage work flow to troubleshoot a certain network error using a machine learning algorithms.

~~The aim of this work is to retrieve key frames and shots from a picture collection, using an algorithm of Sivic and Zisserman [9]. The algorithm employs methods from statistical text retrieval, including inverted file systems, text and document frequency weightings. The advantage of using text retrieval techniques is that the object can be retrieved with no delay due to massive groundwork.~~

The object instances are represented by regions that differ from their environment in their intensity, called **maximally stable extremal regions (MSER)**. Each region is then represented as a vector using the **scale invariant feature transform (SIFT)** algorithm, which efficiently detects and describes local features in images.

Those vectors will compound a database, which will be partitioned into clusters. Each cluster is a **visual word** representing various appearances of the same object in various pictures . Thereafter we remove the most frequent visual words that occur in almost all images, (**“stop words”**)**.** Next, for every visual word, we calculate the **term frequency–inverse document frequency (tf–idf)** index, which is a numeric statistic that reflects the significance of a visual word in characterizing a photo, relatively to other photos in the collection. The last step in the pre-processing is building an **inverted file indexing structure**, which lists the images in which every visual word appears. All that is described above happens at pre-processing time.

At runtime, the user selects a part of a picture that they want to retrieve - the query region. The query region will go through the same steps as the pictures in the dataset go through during the pre-processing. Then, we scan the database and find the most suitable vectors that match to the query vectors. Finally, we search for the key frames that match the set of the query frame vectors.