Case Study

Zuccheti s.p.a

always looking beyond!

https://www.zucchetti.it/

# Description

Zucchetti offers a range of products in Italy and Europe, allowing customers to gain major competitive advantages and to rely on a single partner for all their IT needs. Software and hardware solutions, innovative services designed and developed to meet the specific needs of (i) companies in any sector and of any size, including banks and insurance companies; (ii) professionals (public accountants, labour consultants, lawyers, bankruptcy liquidators, public notaries, etc.), classified associations and tax advisory centres; (iii) local and central public administrations (municipalities, provinces, regions, ministries, public corporations, etc.).

A multi-disciplinary approach is used to identify the needs of customers and to coordinate, manage and create projects, which allows the Zucchetti Group to develop products and services of the highest quality. To get this goal continuous monitoring of the performances of Zucchetti IT infrastructure is required.

Some key applications are based on an infrastructure including three main components: an Nginx frontend dispatches calls to a Tomcats Cluster that directs queries to an SQL Database. These three components generate separated log files that must be unified to develop a consistent performance analysis. Relating Nginx and Tomcat is quite simple because time, IP address, and the URL of the calls identify unique cases. Relating Tomcat and the SQL Database is more complex because Tomcat aggregates connections for READ operations in pools (WRITE operations have a reserved connection), moreover the SQL log is truncated, excluding queries with low execution time.

Performance analysis is aimed at detecting the overall performance of the system, to be aware of performance drops the users may have experienced in specific periods. To this aim outlier cases should de be excluded from the analysis as they can significantly shift the statistics without having a real impact on the system performance experienced by the users. Variant analysis can support the identification of the segment (group of variants) that capture the common behaviour of the system and that can be obtained by verifying which percentile is below a performance score that is considered safe.

Performance analysis is also aimed at identifying the possible causes of performance drops. To this aim variants that are strongly correlated to performance drops must be isolated to assess if significant dependences with performance drops can be verified.

# Assignment

1. Specify the goals of Zucchetti and define a Knowledge Uplift Model that can support the company in reaching them.
2. Unify the Nginx and Tomcat log files to create an event log.
3. Test strategies to unify the Tomcat and the SQL log files.
4. Identify the most significant performance drops.
5. Using variant analysis verify the percentile that is below a safe performance score and how it varies over time.
6. Verify if dependencies between specific variants and performance drops can be considered statistically significant.
7. Based on the results obtained study if control-flow patterns could be used to identify performance drops and if you think predictive analytics could be tested.

# Dataset

The dataset provided by Zucchetti is available in the Files section of the BIS MS Teams.