

1- Compreensão do Negócio

Identificar os especialistas na organização: Allan Vital, Henrique Campos e Kosh Narek.

Levantar e esboçar as necessidades e expectativas: A empresa parceira possui uma plataforma que 9 a cada 10 brasileiros acessam por dia, o que demanda grande disponibilidade, para garantir a eficiência das suas aplicações a empresa parceira deseja realizar a mitigação do risco de indisponibilidade dos sistemas, através de um modelo de IA, que irá prever possíveis falhas na aplicação e emitirá alertas, com isso, haverá melhora na experiência de seus usuários, garantindo a manutenção do número de usuários em sua plataforma, influenciando em suas regras de negócio, e mantendo sua rentabilidade.

Inicialmente serão coletados os dados da aplicação teste, o monitoramento será realizado através da ferramenta Prometheus, estes registros serão salvos no banco de dados do próprio Prometheus, eles serão utilizados para treinar o modelo de IA, do modelo supervisionado.

O modelo de IA será responsável por fazer a predição de possíveis indisponibilidades do sistema e emitir alertas.

Para uma boa volumetria de dados, será utilizado a ferramenta Locust para gerar o estresse da aplicação teste.

Levantamento dos hardwares e softwares: Prometheus, Grafana, Python, Spring boot, Vue, PostgreSQL, Docker Container e AlertManager.

2- Compreensão dos Dados

Conhecer e entender os dados disponíveis:

- Métricas de requisições como load, status, conexões, entre outras.
- Recursos de containers tais como CPU, memória, inbound e outbound.
- Recursos de banco de dados, como load de escrita, de leitura, entre outros.

Avaliação da qualidade dos dados disponíveis: Com base nas ferramentas utilizadas para extração das métricas, os dados já possuem uma qualidade compatível para as nossas necessidades, portanto não exigem tratamento adicional.

Verificar se a volumetria dos dados atende ao negócio: As métricas extraídas atendem as expectativas de identificação e previsibilidade de indisponibilidade da aplicação.

3- Preparação dos Dados

Seleção dos dados para análise:

- **container_cpu_load_average_10s:** Value of container cpu load average over the last 10 seconds.
- **container_memory_failcnt:** Number of memory usage hits limits
- **container_fs_io_current:** Number of I/Os currently in progress
- **nginx_vts_server_request_duration_seconds_count:** The total number of observations for: The histogram of request processing time
- **nginx_vts_server_request_seconds:** The average of request processing times in seconds
- **nginx_vts_server_requests_total:** The requests counter
- **nginx_vts_main_connections:** Nginx connections
- **pg_archiver_last_failed_time:** Time of the last failed archival operation
- **pg_db_blk_read_time:** Time spent reading data file blocks by backends in this database, in ms
- **pg_db_blk_write_time:** Time spent writing data file blocks by backends in this database, in ms

Limpeza dos Dados:

Adequar formato dos dados;

Construir novos atributos com base nos atributos existentes (atributos derivados).