# Relatório de Atividades – 18 de setembro

Hoje focamos nos critérios de avaliação, ajustando nosso projeto às exigências do professor. Abaixo, listamos cada um dos critérios, analisando nosso projeto no estágio atual.

## Estrutura e Clareza da Apresentação (15%)

### Situação atual

O grupo já definiu uma estrutura lógica de apresentação: introdução, objetivos, dados, pré-processamento, metodologia, resultados esperados e considerações finais.

### Próximos passos

Produzir os slides de forma alinhada a essa estrutura, garantindo fluidez e clareza narrativa.

## Contextualização e Relevância (10%)

### Situação atual

O projeto está contextualizado no tema Análise de Sentimentos em comentários de vídeos do YouTube, com relevância social, acadêmica e prática.

### Próximos passos

Detalhar melhor a justificativa prática (ex.: como empresas, criadores de conteúdo ou pesquisadores podem usar esses insights).

## Dados e Pré-processamento (15%)

### Situação atual

Já houve início da coleta de dados via API do YouTube. O pré-processamento começou a ser implementado no código (remoção de caracteres especiais, normalização, etc.).

### Próximos passos

Ampliar o dataset, aplicar tokenização, remoção de stopwords, stemming/lemmatização e mostrar exemplos de antes/depois da limpeza.

## Metodologia (20%)

### Situação atual

O grupo definiu o uso de TF-IDF como representação textual inicial e pretende aplicar classificadores supervisionados (Naive Bayes, Regressão Logística).

### Próximos passos

Justificar tecnicamente as escolhas, implementar os modelos, comparar desempenho e, futuramente, testar embeddings (Word2Vec, BERT).

## Resultados e Análise (20%)

### Situação atual

Ainda não existem resultados prontos, apenas previsão de métricas (acurácia, precisão, recall e F1-score) e visualizações da distribuição de sentimentos.

### Próximos passos

Determinar como implementar os modelos, calcular as métricas e produzir gráficos/tabelas que permitam interpretar os resultados de forma clara.

## Discussão Crítica e Limitações (10%)

### Situação atual

O grupo já reconhece limitações como: dataset inicial pequeno, possível desbalanceamento de classes e uso de técnicas básicas.

### Próximos passos

Expandir a base, avaliar alternativas mais robustas (embeddings) e incluir essa reflexão crítica nos resultados futuros.

## Comunicação Visual (5%)

### Situação atual

Ainda não foram preparados slides ou elementos gráficos finais, apenas saídas iniciais em notebook.

### Próximos passos

Criar slides limpos e organizados, com gráficos claros e legendados, seguindo uma paleta consistente.

## Domínio do Tema e Postura (5%)

### Situação atual

O grupo já se organizou para que cada integrante contribua em partes específicas (dados, metodologia, resultados esperados).

### Próximos passos

Treinar a apresentação, garantindo segurança, clareza e equilíbrio na fala de todos os membros.

## Conclusão Parcial

O projeto encontra-se em fase inicial, com avanços em coleta de dados, início do pré-processamento e definição metodológica. Ainda não há resultados consolidados, mas o grupo demonstra clareza sobre os critérios de avaliação do professor e como cada um será contemplado no desenvolvimento. Os próximos passos concentram-se em expandir o dataset, aplicar os modelos definidos e gerar análises com resultados interpretáveis.