**UNIVERSIDADE REGIONAL DE BLUMENAU**

**CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA**

**CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**Disciplina: Processamento de Linguagem Natural**

**Unidade 1 – Introdução ao Processamento de Linguagem Natural (PLN)**

**Data de entrega: 26/08**

**Equipe:** Gabriel Pereira, Gustavo Bozzano, Lucas Gorges e Rodrigo Heinzen

**1. Tema escolhido (Qual tipo de dado textual? Qual o tema? Descrever brevemente):**

Dados de notícias, como classificação por tipo, conteúdo, tempo em alta no portal

**2. Descrever possíveis aplicações dentro de cada tipo de tarefa:**

|  |  |
| --- | --- |
| **Tarefa** | **Descrição** (explicar quais as entradas e saídas esperadas, processos envolvidos, com foco no resultado esperado). |
| Classificação de texto | Classificar por tipo de notícia |
| Extração de informações | Extrair informações sobre os locais mencionados, empresas, figuras públicas e etc |
| Recuperação de informações | Somatória de todas as notícias da semana |
| Chatbot ou agentes | Chatbot com resumo das notícias/temas em alta |
| Outros |  |

**3. Indicar fontes/repositórios de dados relevantes sobre o tema escolhido (sites, textos etc.)**

Notícias na página principal do portal G1 (https://g1.globo.com)

**4. Quais as possíveis dificuldades e obstáculos da aplicação de PLN ao tema escolhido?**

Alteração do formato do site (pouco provável), bloqueio por portal do G1