Trabalho 4
Privacidade Diferencial – Mecanismo Exponencial

Javam Machado

Novembro 2019

1 Objetivo

- Implementar o mecanismo Exponencial, segundo a privacidade diferencial. Realizar consultas sobre uma base de dados através do mecanismo Exponencial implementado e fornecer os resultados de forma correta e privada.

2 Especificação

- Carregue o conjunto de dados "movie_metadata.csv", contendo dados de filmes.
- As consultas deverão respeitar os seguintes $\varepsilon = \{0.1, 1, 10\}$.
- A totalidade do orçamento de privacidade *(privacy budget – ε)* deverá ser respeitada. Portanto, importante saber quando usar **composição paralela ou sequencial**.
- Consultas a serem realizadas sobre a totalidade dos dados:
  1. Q1: Qual o filme com o maior *gross*;
  2. Q2: Qual filme com o maior *gross* para cada *language*;
  3. Q3: Top 3 *country* com a maior quantidade de filmes;
- Assume-se que o conjunto de dados utilizado é o universo de dados possíveis, sendo assim possível calcular a sensibilidade global de forma fácil.

3 Requisitos

- Linguagem: Python
- Meio de entrega: criar um repositório chamado "disciplina_privacidade_2019" no Github e compartilhar com os seguintes e-mails: {andre.luis, iago.chaves, israel.vidal, javam.machado}@lsbd.ufc.br. **Todos os trabalhos da disciplina serão entregues através desse repositório**.

- Criar uma pasta "dp_exponential" no repositório "disciplina_privacidade_2019".
- **As respostas devem estar disponíveis através do "result.csv"**, onde esse arquivo apresenta as seguintes colunas: total_budget, result_q1, result_q2, result_q3, sens_q1, sens_q2, sens_q3. As respostas das consultas para cada budget exigido e suas respectivas sensibilidades deverão estar presentes nesse arquivo.
- Equipes de até 2 pessoas.
- Somente um repositório deve ser criado por equipe.
- O arquivo "README.md" no repositório deve conter os componentes da equipe.

4 Avaliação

- Na avaliação será considerada a:
  1. Privacidade;
  2. Corretude do algoritmo;
  3. Corretude das sensibilidades;
  4. Apresentação da equipe;
  5. Utilidade da resposta;
  6. Composição de forma adequada;
  7. Qualidade e legibilidade do código.

5 Entrega

- 21 de Novembro de 2019. *Commit* do código até 13:59.