

Relatório Sprint 0

Resumo

Este relatório descreve como se deu o andamento das atividades desenvolvida pela equipe durante a sprint 01.

Mock

O objetivo desta atividade foi desenvolver um gerador de dados, de forma aleatória, para simular os dados reais, tendo em vista que os dados reais ainda não estão disponíveis.

Para tanto, foi desenvolvido um script que, a partir de uma matriz de transição, gera um JSON com as amostragens que simula o uso do produto pelos usuários.

As matrizes de transições foram construídas de duas maneiras: a partir de grafos construídos a mão; e outras aleatoriamente através da distribuição de dirichlet.

As amostragens (sequências de telas de cada usuário) foram geradas utilizando a distribuição multinomial nas probabilidades condicionais de cada tela, ou seja, dado que o usuário se encontra na tela *A*, a escolha da próxima tela é dada a partir de um valor aleatório gerado a partir da distribuição multinomial das probabilidades condicionais de *A* para as demais telas.

Dado que o JSON é um array contendo as atividades de cada usuário, o formato do JSON se dá da seguinte maneira:

```
[
  {
    "IDUSER": 1,
    "IDOWNER": 1,
    "IDPRODUCT": 2,
    "IDFRAME": 1,
    "TIMESTAMP": "1525172983"
  },
  ...
  {
    "IDUSER": X,
    "IDOWNER": 1,
    "IDPRODUCT": 2,
    "IDFRAME": N,
```

```
        "TIMESTAMP": "T"  
    },  
]
```

de modo que X é um usuário qualquer e N é a tela que o usuário se encontra no tempo T.

Avaliação do Protótipo

Para avaliação do protótipo foram utilizadas duas bases de dados, sendo estas da seguinte maneira:

- a primeira foi baseada em três grafos criados a mão, gerando 3 matrizes de transições 30x30, representando assim 1 produto e 3 usuários
- a segunda foi feita a partir de matrizes transições 30x30 geradas aleatoriamente, representando 3 produtos e 10 usuários.

As amostragens geradas foram separadas em treino e teste, nas quais, a partir das amostragens de treino, gerou-se uma matriz de transição baseada nos dados de treinamento e com as amostragens de teste foi verificado se o protótipo foi capaz de sugerir corretamente a próxima tela utilizando a matriz de transição encontrada no treinamento.

Um parâmetro de *rank* definiu a quantidade de sugestões que o protótipo recomendaria, assim, se a recomendação correta estivesse entre as recomendações do rank, seria considerado um acerto durante o teste.

As Tabela 1 e 2 exibem os resultados obtidos com as bases de dados utilizadas.

<i>Rank</i>	Média Usuário	Média Produto
1	0.7195	0.1956
2	0.9149	0.4172
3	0.9709	0.6940
4	0.9944	0.8158
5	1.0000	0.9428

Tabela 1: Resultados para a base de dados gerada a partir dos grafos criados

<i>Rank</i>	Média Usuário	Média Produto
1	0.4548	0.0272
2	0.6522	0.0619
3	0.7619	0.1013
4	0.8384	0.1209
5	0.8878	0.1783

Tabela 2: Resultados para a base de dados gerada aleatoriamente