

# Recommender Systems Lab2

**Due Date: 11:59PM May 11, 2024**

You will have a UN201301.txt file. Each row in the UN201301.txt (encoding is UTF-8) denotes a news story from the UDN.com. The meanings of columns are introduced as follows.

- The 01st column: PaperID
- The 02nd column: PresentDate
- The 03rd column: HotLine
- The 04th column: SubHotLine
- The 05th column: Author
- The 06th column: Head
- The 07th column: Head\_Name
- The 08th column: PageName
- The 09th column: PageNo
- The 10th column: Document

**Q1: Which date has the most news stories?**

**Q2: How many distinct hotlines?**

**Q3: How many sentences (separated by a period) in all news documents?**

**Example outputs**

```
Q1: Which date has the most news stories? {your ans}
Q2: How many distinct hotlines? {your ans}
Q3: How many sentences (separated by a period) in all news documents? {your ans}
```

**Hint:**

1. Use split function
2. How to load txt file

```
f = open('UN201301.txt', 'r', encoding="utf-8")
for line in f.readlines():
    print(line)
f.close()
```