

Text Summarization

Nguyễn Đình Hoàng Phúc
Nguyễn Mai Kiều Anh
Hoàng Như Ngọc

MaSSP 2018

Ngày 3 tháng 7 năm 2018

Outline

Problem description

Applications

Machine learning model

Problem description

Mô tả vấn đề và ý tưởng giải quyết

► 500.000.000

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ **500.000.000**

Số lượng báo giấy (newspaper) bán ra trong một ngày

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ **500.000.000**
Số lượng báo giấy (newspaper) bán ra trong một ngày
- ▶ **5.500.000**

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ **500.000.000**
Số lượng báo giấy (newspaper) bán ra trong một ngày
- ▶ **5.500.000**
Số lượng bài blog được viết trong một ngày

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ **500.000.000**
Số lượng báo giấy (newspaper) bán ra trong một ngày
- ▶ **5.500.000**
Số lượng bài blog được viết trong một ngày
- ▶ **2.200.000**

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ **500.000.000**
Số lượng báo giấy (newspaper) bán ra trong một ngày
- ▶ **5.500.000**
Số lượng bài blog được viết trong một ngày
- ▶ **2.200.000**
Số lượng sách xuất bản trong một năm

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ Vấn đề:
 - Lượng dữ liệu dạng văn bản rất lớn và đang tăng lên
 - Không thể trực tiếp giải quyết hết

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ Vấn đề:
 - Lượng dữ liệu dạng văn bản rất lớn và đang tăng lên
 - Không thể trực tiếp giải quyết hết
- ▶ Nhu cầu: Cắt giảm bớt, tóm tắt các chi tiết nổi bật

Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ Vấn đề:
 - Lượng dữ liệu dạng văn bản rất lớn và đang tăng lên
 - Không thể trực tiếp giải quyết hết
- ▶ Nhu cầu: Cắt giảm bớt, tóm tắt các chi tiết nổi bật
- ▶ Ý tưởng giải quyết: Text Summarization
 - **Input:** Văn bản gốc
 - **Output:** Đoạn tóm tắt



Problem description

Mô tả vấn đề và ý tưởng giải quyết

- ▶ Vấn đề:
 - Lượng dữ liệu dạng văn bản rất lớn và đang tăng lên
 - Không thể trực tiếp giải quyết hết
- ▶ Nhu cầu: Cắt giảm bớt, tóm tắt các chi tiết nổi bật
- ▶ Ý tưởng giải quyết: Text Summarization
 - **Input:** Văn bản gốc
 - **Output:** Đoạn tóm tắt



- ▶ Lợi ích:
 - Tiết kiệm thời gian đọc
 - Lưu trữ có hệ thống
 - Tiết kiệm thời gian chọn lọc, tìm kiếm

Outline

Problem description

Applications

Machine learning model

Applications

Ứng dụng và ví dụ thực tế của Text Summarization

- ▶ Tóm tắt tin tức

The Defense Intelligence Agency believes that North Korean leader Kim Jong Un has no intention of engaging in a full denuclearization program, at least for now, according to an administration official familiar with the agency's finding.

→ **US intel agency believes Kim won't fully denuclearize.**

Applications

Ứng dụng và ví dụ thực tế của Text Summarization

- ▶ Tóm tắt tin tức

The Defense Intelligence Agency believes that North Korean leader Kim Jong Un has no intention of engaging in a full denuclearization program, at least for now, according to an administration official familiar with the agency's finding.

→ **US intel agency believes Kim won't fully denuclearize.**

- ▶ Tóm tắt phim, sách

A team of explorers travel through a wormhole in space in an attempt to ensure humanity's survival. (*Interstellar*)

Applications

Ứng dụng và ví dụ thực tế của Text Summarization

- ▶ Tóm tắt tin tức

The Defense Intelligence Agency believes that North Korean leader Kim Jong Un has no intention of engaging in a full denuclearization program, at least for now, according to an administration official familiar with the agency's finding.

→ **US intel agency believes Kim won't fully denuclearize.**

- ▶ Tóm tắt phim, sách

A team of explorers travel through a wormhole in space in an attempt to ensure humanity's survival. (*Interstellar*)

- ▶ Giáo dục: tóm tắt bài giảng/sách giáo khoa/tài liệu

Applications

Ứng dụng và ví dụ thực tế của Text Summarization

- ▶ Tóm tắt tin tức

The Defense Intelligence Agency believes that North Korean leader Kim Jong Un has no intention of engaging in a full denuclearization program, at least for now, according to an administration official familiar with the agency's finding.

→ **US intel agency believes Kim won't fully denuclearize.**

- ▶ Tóm tắt phim, sách

A team of explorers travel through a wormhole in space in an attempt to ensure humanity's survival. (*Interstellar*)

- ▶ Giáo dục: tóm tắt bài giảng/sách giáo khoa/tài liệu

- ▶ etc

Outline

Problem description

Applications

Machine learning model

Machine learning model

Mô hình machine learning sử dụng