



# **B4B33RPH: Řešení problémů a hry**

## **3. úloha: Spam filtr**

Petr Pošík a Tomáš Svoboda



# SPAM?

---

- SPAM?
- Definice
- Spam filtr
- Úloha



# SPAM?

---

- SPAM?
- Definice
- Spam filtr
- Úloha



Hormel Foods Corporation, 1937



# SPAM?

- SPAM?
- Definice
- Spam filtr
- Úloha



Hormel Foods Corporation, 1937

70. léta:

- **Monty Python's Flying Circus** (youtube):
  - Manželé si chtějí dát v bistru snídani.
  - Manželka nechce jídlo obsahující SPAM.
  - Obsluha i manžel jí přesto stále SPAM vnucují.
  - "I DON'T LIKE SPAM!!!"
- **SPAM** - nezajímavé, nechtěné, vnucované, opakující se zprávy
- **HAM** - korektní zprávy týkající se tématu



# Definice?

---

Spam je

- *nevyžádaná* zpráva
- *hromadně rozesílaná* mnoha lidem *bez rozdílu*,
- velmi často *za komerčním účelem*.

- SPAM?
- Definice
- Spam filtr
- Úloha



- SPAM?
- Definice
- Spam filtr
- Úloha

## Definice?

---

Spam je

- *nevyžádaná* zpráva
- *hromadně rozesílaná* mnoha lidem *bez rozdílu*,
- velmi často *za komerčním účelem*.

Zkuste si odpovědět:

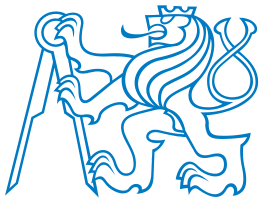
1. Označili by všichni lidé za spam ty samé emaily?
2. Kdyby stejnému člověku přišla stejná sada emailů v současnosti a za pět let, označil by za spamy stejné emaily?

**A** 1. Ne. | 2. Ne.

**B** 1. Ne. | 2. Ano.

**C** 1. Ano. | 2. Ne.

**D** 1. Ano. | 2. Ano.



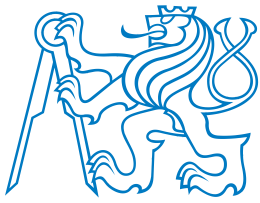
# Spam filtr

---

Spam filtr:

- Automat, který každému emailu umí přiřadit štítek SPAM nebo HAM (přičemž se může plést).

- SPAM?
- Definice
- Spam filtr
- Úloha



- SPAM?
- Definice
- Spam filtr
- Úloha

# Spam filtr

---

Spam filtr:

- Automat, který každému emailu umí přiřadit štítek SPAM nebo HAM (přičemž se může plést).

Na základě čeho se spam filtr rozhoduje?

- A** Podle hlavičky emailu (odesílatel, adresát, datum odeslání, ...)
- B** Podle těla emailu (text zprávy).
- C** Podle počtu, typu a obsahu příloh.
- D** Podle formátu textu (barvy, tučné nebo blikající písmo, ...)





- SPAM?
- Definice
- Spam filtr
- Úloha

# Spam filtr

---

Spam filtr:

- Automat, který každému emailu umí přiřadit štítek SPAM nebo HAM (přičemž se může plést).

Jaké má vstupy a jaké má výstupy?

- Vstup: všechny informace o emailu, které můžeme získat (tělo emailu, přílohy, hlavičky). Velmi často nestrukturovaný text.
- Výstup: štítek SPAM nebo HAM



- SPAM?
- Definice
- Spam filtr
- Úloha

# Spam filtr

---

Spam filtr:

- Automat, který každému emailu umí přiřadit štítek SPAM nebo HAM (příčemž se může plést).

Jaké má vstupy a jaké má výstupy?

- Vstup: všechny informace o emailu, které můžeme získat (tělo emailu, přílohy, hlavičky). Velmi často nestrukturovaný text.
- Výstup: štítek SPAM nebo HAM

Zkuste si odpovědět:

1. Bude jeden a tentýž spam filtr vyhovovat všem lidem?
2. Co musí spam filtr umět, aby si udržel svou efektivitu v průběhu času?
3. Co musí umět, aby byl efektivní pro různé lidi?
4. Může se filtr dopustit chyb? Jakých?
5. Budou všechny typy chyb stejně vážné?
6. Jak poznám, že je jeden filtr lepší než druhý?



- SPAM?
- Definice
- Spam filtr
- Úloha

# Úloha

---

Proč byl spam filtr vybrán jako úloha v předmětu RPH?

- Praktická netriviální úloha, přesto řešitelná s našimi znalostmi Pythonu.
- Prohloubení schopnosti zpracovat text, pracovat se soubory a datovými strukturami.
- Ukázka učení z dat (strojového učení).

Na práci na spam filtru máme cca 4 týdny.

- 2 týdny: **individuální** práce na “zázemí”, které nám umožní filtry
  - jednoduše psát,
  - snadno testovat a porovnávat a
  - vyjasnit si, jak že vlastně ten spam filtr bude vypadat.
- 2 týdny: **týmová** práce na vlastním filtru.
  - Vyzkoušíte si vámi zvolenou metodu filtrování spamu.

Dostanete testy k jednotlivým dílčím podúlohám.

Odevzdávat budete

- implementaci funkce pro hodnocení kvality filtru (každý sám),
- implementaci vašeho spam filtru (ve dvojicích) a
- report + prezentaci (ve dvojicích, bude doupřesněno).

Bodování: nápad, čistota kódu, účinnost na skutečných datech



# Díky za pozornost

---

- SPAM?
- Definice
- Spam filtr
- Úloha

Dotazy?