

# Sistem za prepoznavanje lažnih vijesti

## Članovi tima

- Bojan Živanić SV61/2021

## 1. Opis problema

### Motivacija

U digitalnom društvu veliki dio informacija dolazi sa interneta i društvenih mreža. Brzo širenje neprovjerenih i lažnih vijesti utiče na javno mnjenje, obrazovanje i donošenje odluka građana. Cilj projekta je razvoj ekspertnog sistema koji omogućava korisnicima da identifikuju potencijalno lažne vijesti koristeći pravila, transparentno rezonovanje i analizu obrazaca u sadržaju i izvorima.

### Pregled problema

Većina postojećih rješenja koristi metode mašinskog učenja i NLP, što zahtijeva velike datasetove i složene modele. Takvi sistemi često nisu transparentni.

Predloženi sistem je rule-based, što omogućava:

- Vidljivost i objašnjenje odluka sistema
- Korišćenje forward i backward chaininga
- Analizu ponavljanja vijesti kroz CEP
- Standardizovan unos podataka putem templejta

Prednost sistema je jednostavna implementacija, transparentno rezonovanje i lako prilagođavanje novim pravilima i izvorima.

## 2. Metodologija rada

### Ulazi (Input)

- Izvor vijesti (poznat, nepouzdan, nepoznat)
- Datum objave
- Naslov i sadržaj (ključne riječi, stil)
- Broj ponavljanja iste vijesti u kratkom vremenskom periodu
- Kontekst (drugi portali gde je vijest objavljena)

## Izlazi (Output)

- Kategorija pouzdanosti:
  - Pouzdana vijest
  - Potencijalno lažna vijest
  - Sumnjiva vijest – potrebno dodatno proveriti
- Obrazloženje odluke sistema

## Baza znanja

- Pravila o reputaciji izvora: poznat, nepouzdan, nepoznat
- Pravila o sadržaju: senzacionalističke fraze, pretjerani uzvičnici, netačne reference
- Pravila o ponavljanju vijesti (CEP)
- Template za uniforman unos vijesti, uključujući izvor, datum, stil i sadržaj

## 3. Forward chaining

Sistem koristi pravila koja se ulančavaju u tri nivoa:

### Nivo 1 – Reputacija izvora

- Ako je izvor u listi poznatih portala, reputacija = *visoka*
- Ako je izvor u listi nepouzdatih portala ili je nepoznat, reputacija = *niska*

### Nivo 2 – Analiza sadržaja

- Ako je reputacija *visoka* i nema senzacionalizma, preliminarna ocjena = *pouzdana vijest*
- Ako je reputacija *visoka* i ima senzacionalizam, preliminarna ocjene = *potencijalno lažna vijest*
- Ako je reputacija *niska* i nema senzacionalizma, preliminarna ocjena = *potencijalno lažna vijest*
- Ako je reputacija *niska* i ima senzacionalizam, preliminarna ocjena = *sumnjiva vijest*

### Nivo 3 – CEP / Ponavljanja

1. Ako je preliminarna ocjena *pouzdana vijest* i ponavljanje  $\leq 3$  puta, finalna ocjena = *pouzdana vijest*
2. Ako je preliminarna ocjena *pouzdana vijest* i ponavljanje  $> 3$  puta na nepouzdatim portalima, finalna ocjena = *sumnjiva vijest*
3. Ako je preliminarna ocjena *potencijalno lažna vijest* i ponavljanje  $\leq 3$  puta, finalna ocjena = *potencijalno lažna vijest*
4. Ako je preliminarna ocjena *potencijalno lažna vijest* i ponavljanje  $> 3$  puta na nepouzdatim portalima, finalna ocjena = *sumnjiva vijest*
5. Ako je preliminarna ocjena *sumnjiva vijest*, finalna ocjena = *sumnjiva vijest* (bez obzira na broj ponavljanja)

## 4. Backward chaining

Zasniva se na postavljanju cilja i dokazivanju uslova unazad. Sistem kreće od pitanja: *Da li je vijest pouzdana?*

Koraci rada sistema:

- Postavljanje cilja - sistem želi da potvrdi ili opovrgne hipotezu da je vijest pouzdana
- Provjera uslova - da bi vijest bila pouzdana, moraju biti ispunjeni sljedeći uslovi:
  - reputacija izvora = visoka
  - nema senzacionalizam u naslovu i sadržaju
  - ponavljanje  $\leq 3$  puta
- Postavljanje podciljeva - ako neki uslov nije poznat, sistem ga dalje razlaže:
  - za reputaciju provjerava listu pouzdanih/nepouzdanih izvora
  - za sadržaj analizira stil i ključne riječi
  - za ponavljanje koristi CEP modul da prebroji pojavljivanja vijesti
- Izvođenje zaključka - ako su uslovi ispunjeni, vijest je pouzdana. Ako neki uslov nije ispunjen, sistem pokušava alternativne hipoteze:
  - potencijalno lažna vijest
  - sumnjiva vijest
- Finalna odluka - sistem daje kategoriju i obrazloženje zasnovano na tragovima rezonovanja

Na ovaj način backward chaining omogućava da korisnik jasno vidi zašto vijesti nije prošla određeni kriterijum i razlog zbog kog je svrstana u određenu kategoriju.

## 5. CEP (Complex Event Processing)

CEP modul koristi se za analizu obrazaca pojavljivanja vijesti u realnom vremenu.

Pravila rada CEP-a:

- Sistem prati koliko puta se ista vijest pojavljuje u određenom vremenskom periodu
- Ako je broj ponavljanja  $\leq 3$  puta, vijest se smatra normalno distribuiranom
- Ako je broj ponavljanja  $> 3$  puta, posebno na nepouzdanim portalima, vijest se označava kao sumnjiva.

Prednost CEP-a je u tome što otkriva obrasce širenja lažnih vijesti koje se često pojavljuju koordinisano i intenzivno. Na taj način sistem ne analizira samo sadržaj, već i dinamiku distribucije vijesti.

## 6. Primjer konkretnog rezonovanja

### Unos vijesti:

- Izvor: Nepoznat portal
- Naslov: „ŠOKANTNO! Ljekari sakrili istinu o vodi koju pijemo!“
- Datum: Juče
- Ponavljanja: 5 puta

### Koraci rezonovanja:

1. Sistem analizira izvor vijesti i utvrđuje nisku reputaciju.
2. Analizira se naslov i sadržaj vijesti, gdje se prepoznaju senzacionalističke fraze.
3. CEP funkcionalnost prepoznaje da se vijest pojavila više puta u kratkom vremenskom periodu na nepouzdanim portalima.
4. Backward chaining pokušava dokazati hipotezu *vijest je pouzdana*, ali uslovi nisu ispunjeni:
  - reputacija nije visoka
  - naslov je senzacionalistički
  - ponavljanje > 3 puta

Hipoteza se odbacuje, sistem prelazi na provjeru alternativnih hipoteza

5. Na osnovu pravila, sistem zaključuje da vijest spada u kategoriju *sumnjivih vijesti*

### Izlaz sistema:

- Kategorija: *Sumnjiva vijest*
- Obrazloženje: Vijest dolazi sa nepoznatog izvora, ima senzacionalistički naslov i višestruko se pojavila na nepouzdanim portalima.