

## Fișă proiect instituțional pentru publicarea pe site-ul cercetare.ase.ro

(maxim 5 pagini)

### Titlu proiect:

**MODELAREA MATEMATICĂ A FACTORILOR CARE DUC LA ABONAREA/DEZABONAREA DIN LISTE DE EMAIL/SMS, FORUMURI, GRUPURI DE SOCIAL MEDIA**

### Instituția care a propus tema:

**ACADEMIA DE STUDII ECONOMICE BUCUREȘTI – FACULTATEA DE CIBERNETICĂ, STATISTICĂ ȘI INFORMATICĂ ECONOMICĂ**

### Director:

**Prof. dr. Dârdală (Reveiu) Adriana Elena**

### Echipa:

**Prof. univ. dr. Agapie Alexandru Licinius  
Conf. univ. dr. Covrig Mihaela  
Drd. Arghir Denis-Cătălin  
Drd. Vîlcea Lucian-Alexandru**

### Obiective:

Obiectivul principal al proiectului este identificarea factorilor care contribuie la schimbarea deciziilor consumatorilor principalelor medii de comunicare online: liste de distribuție, grupurile de email și rețele de socializare, în vederea modelării matematice a comportamentului consumatorilor în mediul digital.

În vederea modelării matematice a comportamentului utilizatorilor în diverse medii digitale, cercetarea a propus răspunsuri la următoarele întrebări, ce reprezintă totodată obiectivele secundare:

Q1: Care sunt factorii ce influențează decizia consumatorilor de a își schimba comportamentul în mediul online? Care este importanța acestora în influențarea deciziilor?

Q2: Care sunt tehnicile de modelare care oferă cele mai bune estimări ale contribuției factorilor în previzionarea comportamentului consumatorilor, în fiecare context?

### Activități:

Activitățile derulate în cadrul proiectului sunt:

## **1. Achiziția datelor relevante pentru modelarea comportamentului consumatorilor în mediul digital**

1.1 Identificarea surselor de date relevante pentru diverse tipuri de comunități digitale: liste de distribuție, grupuri publice, rețele de socializare

1.2 Identificarea și analiza tehnologiilor relevante pentru accesul și achiziția datelor

1.3 Realizarea componentelor software pentru preluarea datelor din mediul online

## **2. Procesarea și analiza datelor semi-structurate și structurate**

2.1 Automatizarea preluării metadatelor relevante referitoare la resursele de tip imagine

2.2 Modelarea și stocarea datelor achiziționate într-o bază de date non-relațională

## **3. Modelarea matematică și analiza comportamentului consumatorilor în mediul digital**

3.1 Modelare bazată pe analiza sentimentelor a mesajelor text preluate din listele de distribuție și rețele de socializare

3.2 Adaptarea tehnicilor de *machine learning* în vederea clasificării consumatorilor prin identificarea factorilor potențiali

3.3 Modelarea folosind lanțuri Markov a comportamentului consumatorilor în mediul digital

3.4 Testarea modelelor pe diverse seturi de date

## **4. Diseminarea și prezentarea rezultatelor cercetării**

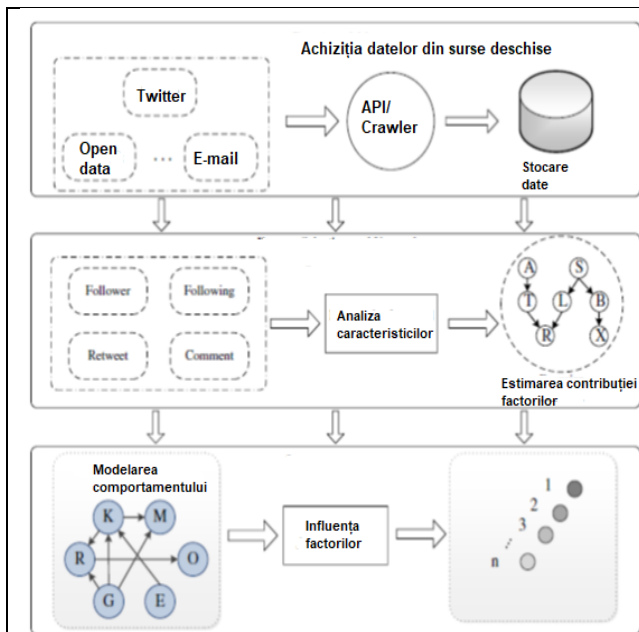
### **Rezultate:**

Dezvoltarea tehnologiilor digitale a schimbat fundamental modul de interacțiune al companiilor cu clienții și potențialii clienți, consumatorii obișnuind să comunice prin intermediul rețelelor online experiențele avute cu diverse companii, indiferente dacă acestea au fost pozitive sau negative. În această manieră, opiniile clienții ajung rapid la un număr important de persoane interesate, iar părerile nefavorabile pot genera pierderile materiale și de imagine semnificative pentru companii. Apariția tehnologiilor *Big data* a permis gestiunea și prelucrarea volumelor mari de date vehiculate în Internet, precum cele postate pe rețelele de socializare. Acestea s-au dovedit relevante pentru analiza comportamentului și sentimentului consumatorilor, a impactului situațiilor de criză asupra imaginii brandului, prognoza unor indicatori macro-economici și pentru identificarea unor idei de afaceri.

În acest context, proiectul propune noi soluții de identificare și modelare a factorilor care contribuie la fundamentarea și schimbarea deciziilor consumatorilor în mediul online. Metodologia de cercetare construită în vederea atingerii obiectivelor proiectului s-a bazat pe explorarea tehnicilor de *machine learning* și modele stohastice, adaptate în vederea automatizării prelucrării volumelor mari de date, disponibile în mediul online.

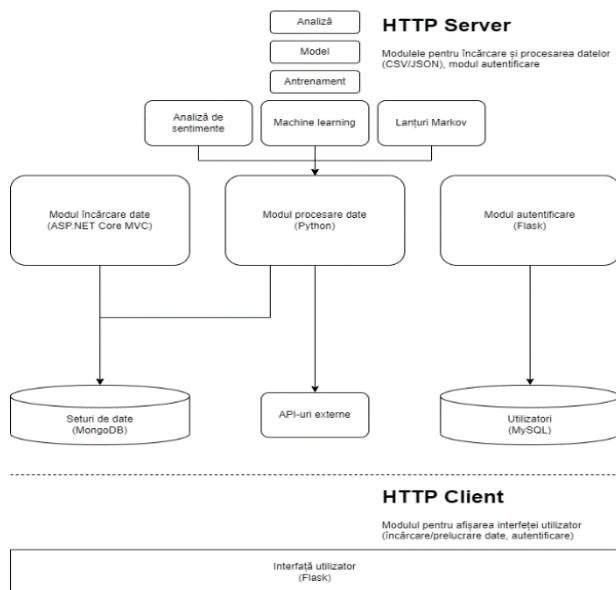
Pentru a asigura flexibilitate în utilizare și dezvoltarea ulterioară a soluției, precum și pentru a facilita tuturor companiilor și instituțiilor interesate să folosească *framework*-ul pentru prelucrarea propriilor seturi de date, s-a dezvoltat un prototip de platformă software ce urmărește utilizarea de soluții *open source* pentru date și pentru dezvoltarea componentelor software de prelucrare și modelare.

*Framework*-ul propus pentru achiziția, prelucrarea și modelarea comportamentului consumatorilor în mediul digital și a factorilor ce influențează schimbarea deciziei este reprezentat în schematic în figura 1.



în medii online (figura 4),

**Figura 1 - Framework pentru modelarea comportamentului consumatorilor**



**Figura 2 - Arhitectura platformei software pentru modelarea comportamentului consumatorilor**

Arhitectura platformei software proiectată și implementată ca soluție prototip integrează o serie de componente software dezvoltate pentru automatizarea prelucrării și modelării datelor și este prezentată în figura 2.

Componentele sistemului software prototip (figura 3), dezvoltate și utilizate în cadrul proiectului sunt folosite pentru:

- Achiziția datelor din mediul online, automatizată pentru rețeaua de socializare Twitter.
- Analiza exploratorie, în vederea identificării metodelor relevante pentru fiecare mediu online.
- Modelarea folosind tehnici de *machine learning* pentru modelarea factorilor ce influențează comportamentul utilizatorilor

- Modelarea stohastică a comportamentului diverselor categorii de consumatori.
- Aplicarea tehnicilor de prelucrare a limbajului natural, pentru extragerea sentimentelor consumatorilor și integrarea acestora în model.
- Analiza performanțelor și acurateței modelelor aplicate, în vederea identificării modelului optim.

Pentru stocarea și gestiunea datelor preluate din diverse surse și pentru automatizarea procesului de gestiune și prelucrare a seturilor de date utile modelării comportamentului consumatorului a fost proiectată și utilizată o bază de date MongoDB, iar datele stocate în format NoSQL.

Rezultatele proiectului au potențialul de a avea un impact semnificativ atât din punct de vedere teoretic, cât și practic. Din perspectivă

practică, proiectul oferă un set de modele fundamentate matematic, un *framework* de analiză, prelucrare și modelare a datelor, o platformă software pilot dezvoltată folosind soluții *open source* - suport pentru analiza diferitelor seturi de date, precum și o serie de studii de caz utile atât pentru dezvoltarea activității de cercetare, cât și companiilor și instituțiilor interesate de investigarea comportamentului utilizatorilor în mediul online.

Parte dintre rezultatele obținute în cadrul proiectului au fost diseminate în lucrări științifice publicate în jurnale indexate în *Web of Science*.



Figura 3 - Interfața principală a aplicației web

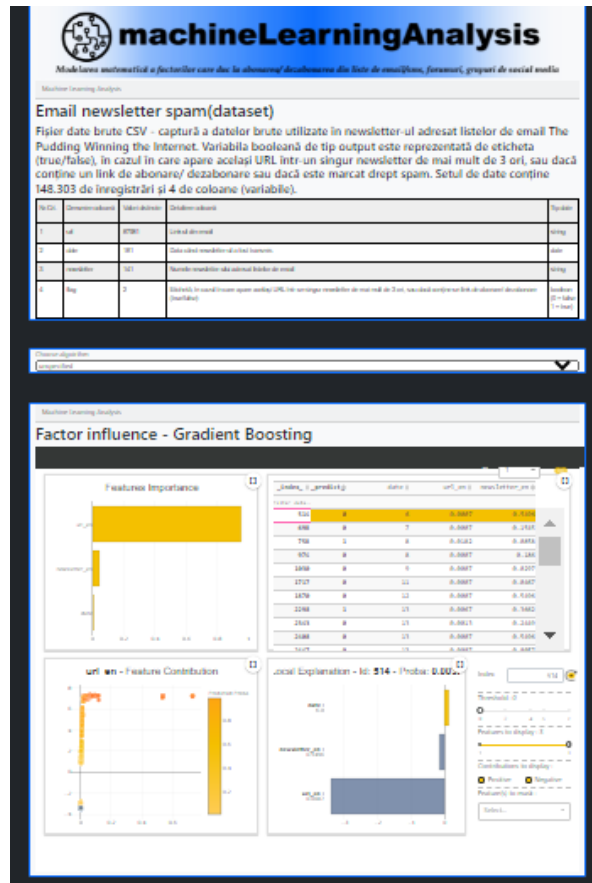


Figura 4 - Interfața componentei de analiză și modelare a setului de date ref. la liste de distribuție