

# Instagram Analytics

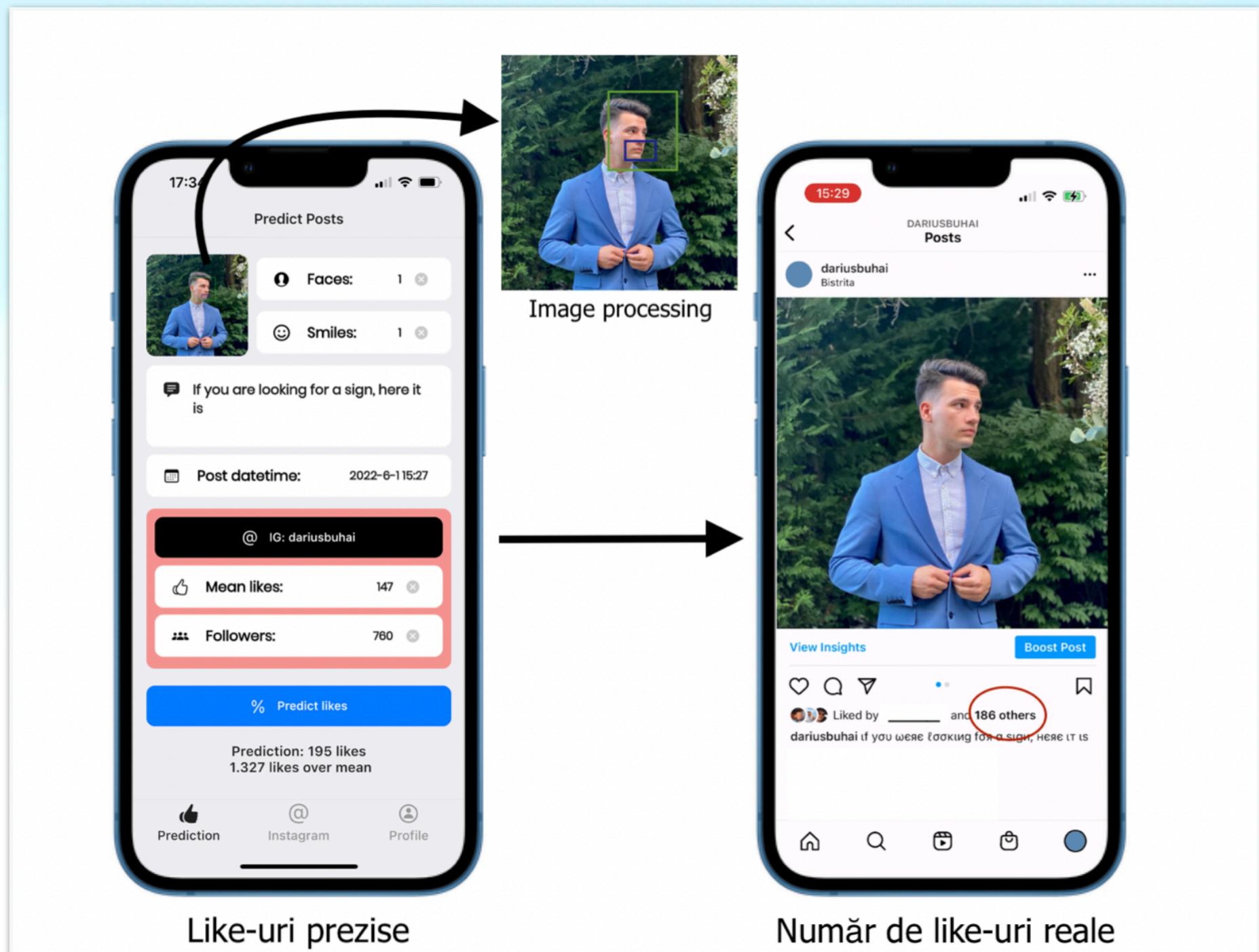
Lucrare de Licență

Proiect realizat de: **Buhai Darius**  
Profesor coordonator: **Dumitran Adrian-Marius**



# Motivăție

- Creșterea numărului de like-uri primite în postările de Instagram, cât și analiza asupra urmăritorilor activi reprezintă niște ustensile destul de importante pentru orice influencer de success.



# Preluarea Datelor

## Conturile de instagram

- Pentru a prelua postările de Instagram, vom avea nevoie mai întâi de influenceri de la care să extragem aceste postări

A Country	# Rank	A Account	A Title	↳ Link	A Category	# Followers	A Audience ...
All	1	cristiano	Cristiano Ronaldo	<a href="https://www.instagram.com/cristiano/">https://www.instagram.com/cristiano/</a>	Sports with a ball	400100000.0	India
All	2	kyliejenner	Kylie 💕	<a href="https://www.instagram.com/kyliejenner/">https://www.instagram.com/kyliejenner/</a>	Fashion Modeling Beauty	308800000.0	United States
All	3	leomessi	Leo Messi	<a href="https://www.instagram.com/leomessi/">https://www.instagram.com/leomessi/</a>	Sports with a ball Family	306300000.0	Argentina
All	4	kendalljenner	Kendall	<a href="https://www.instagram.com/kendalljenner/">https://www.instagram.com/kendalljenner/</a>	Modeling Fashion	217800000.0	United States
All	5	selenagomez	Selena Gomez	<a href="https://www.instagram.com/selenagomez/">https://www.instagram.com/selenagomez/</a>	Music Lifestyle	295800000.0	United States
All	6	zendaya	Zendaya	<a href="https://www.instagram.com/zendaya/">https://www.instagram.com/zendaya/</a>	Cinema Actors/actresses Fashion	127800000.0	United States
All	7	kimkardashian	Kim Kardashian West	<a href="https://www.instagram.com/kimkardashian/">https://www.instagram.com/kimkardashian/</a>	Fashion Beauty	284900000.0	United States
All	8	beyonce	Beyoncé	<a href="https://www.instagram.com/beyonce/">https://www.instagram.com/beyonce/</a>	Music Fashion	237200000.0	United States

# Preluarea Datelor

## Postările

- instagram\_private\_api
- Facebook Graph API

username	followers	engagement	description	taken_at	image_url
cristiano	400100000.0	7800000.0	<null>	1651597276.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Once again, great support from the stands. Let's take this opportunit...	1651528572.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Forever Love...❤️	1651312178.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Thanks for the support in Old Trafford tonight. As always, our fans w...	1651179311.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	♥	1650730663.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Home sweet home. Gio and our baby girl are finally together with us. ...	1650571036.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	One world... One sport... One global family... Thanks, Anfield. Me and my f...	1650536211.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	É com a mais profunda tristeza que comunicamos o falecimento do nosso bebé. É a maior dor que quaisquer pais podem sentir. Só o nascimento da nossa...		
cristiano	400100000.0	7800000.0	30 hat-tricks before 30 and 30 hat-tricks after 30. It's time to unba...	1650205457.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Very happy with this win and for getting us back on track in the Prem...	1650126414.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	⚽️♂ #mufc	1649872077.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	It's never easy to deal with emotions in difficult moments such as th...	1649537161.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Work hard, focus on the everyday goals, enjoy the moments with your f...	1649362119.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	A club such as ours is not only measured by the number of trophies in...	1649329441.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Work done 🤘	1649260991.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Individual achievements are always welcome, but they're worth so much...	1649177336.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Family is everything ❤️ #proud	1648991801.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	You need more than quick feet to be a champion. It takes hours of har...	1648745970.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Há muitas formas de alcançar o destino desejado, por vezes o caminho ...	1648657412.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Os anos passam, mas os objetivos continuam sempre os mesmos! 🇵🇹	1648632108.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Objectivo atingido, estamos no Mundial do Catar, estamos no nosso dev...	1648590598.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Uma nação valente e imortal! Todos juntos e a uma só voz! Força Portu...	1648494405.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	🇵🇹❤️	1648312545.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Had a fun day with @livescore Download the LiveScore app to see wha...	1648285269.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Recovery 🇵🇹	1648234633.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Está dado o primeiro passo rumo ao nosso grande objectivo, rumo ao Mu...	1648162512.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Não é segredo para ninguém: representar o meu país é sempre um moment...	1648059367.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Foco total no Mundial 2022. Orgulhoso, como sempre, por representar P...	1647899015.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	We're opening a new Insparya clinic in Valencia! The 3rd in Spain and...	1647856050.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Happy father's day to all! ❤️ #prouddad #blessed #love	1647706514.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Tomorrow is one of those days that we've all been waiting for. One of...	1647279445.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>
cristiano	400100000.0	7800000.0	Never let doubts or dandruff hold you back. When the time comes, cle...	1647263785.0	<a href="https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2">https://instagram.fotp3-1.fna.fbcdn.net/v/t51.2885-15/2</a>

# Extragerea caracteristicilor

## Din imagini

- Folosim OpenCV pentru a extrage caracteristicile din imagini (numarul de zambete si fete)



# Extragerea caracteristicilor

## Din descrieri

- Din vocabularul nostru putem observa următoarele 6 cele mai frecvente cuvinte:
  1. '**thi**' (35,054 de utilizări), acest token reprezintă cuvântul 'this'.
  2. '**love**' (23,997 de utilizări). Forme incluse: 'loving', 'loved', 'lover' etc.
  3. '**thank**' (18,295 de utilizări). Forme incluse: 'thanks', 'thank you' etc.
  4. '**happi**' (11,380 de utilizări). Forme incluse: 'happy', 'happier' etc
  5. '**time**' (10,395 de utilizări). Forme incluse: 'times', 'timer' etc.
  6. '**link**' (9,338 de utilizări). Forme incluse: 'links' etc.

# Extragerea caracteristicilor

## Alte informații

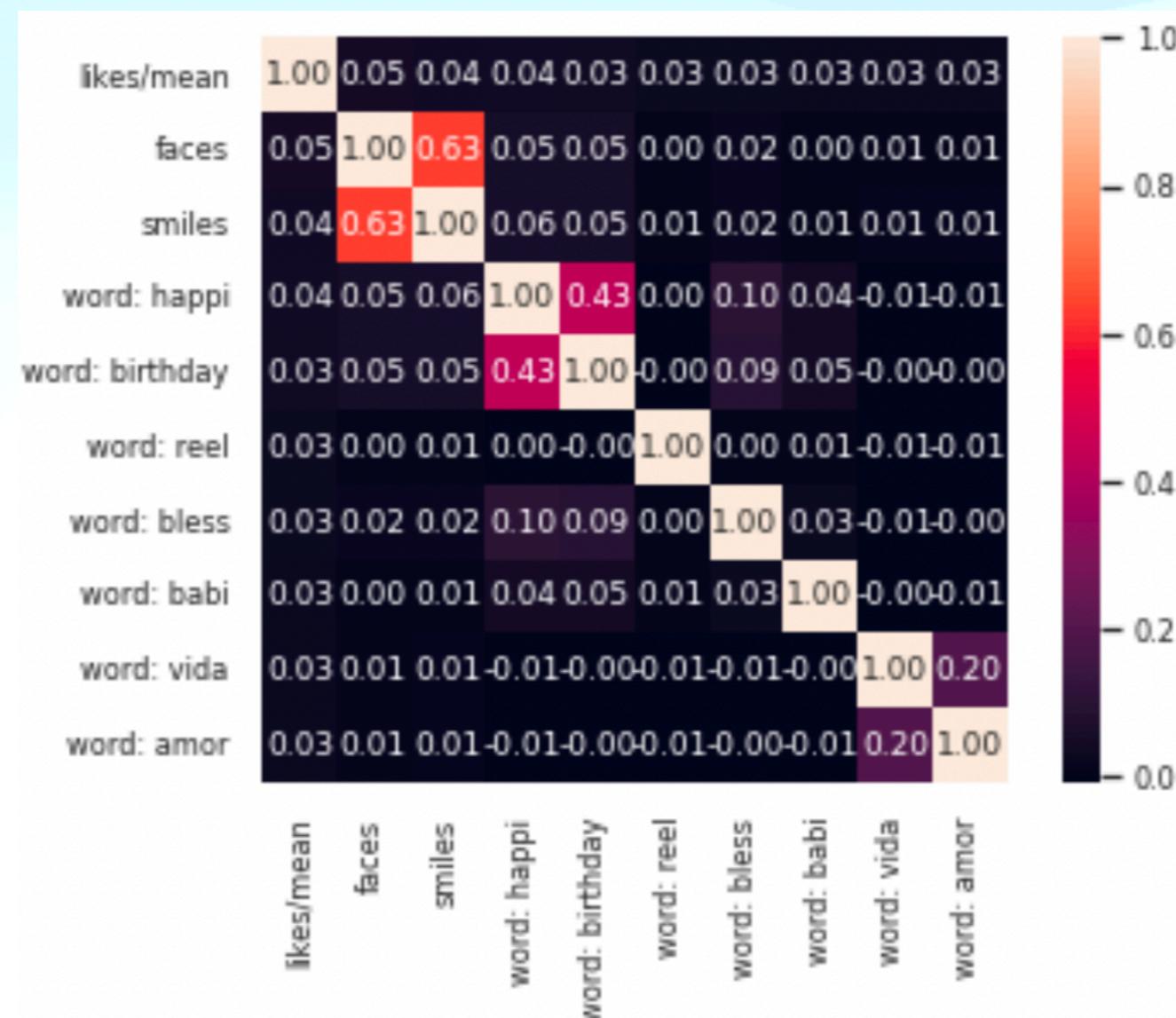
- Astfel, pentru fiecare postare în parte am adăugat 13 caracteristici noi, reprezentate de ziua săptămânii în care a fost postată (Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday) și intervalul orar (8:00-12:00, 12:00-17:00, 17:00-22:00, 22:00-2:00, 2:00- 8:00).

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	
<null>	smiles	faces	followers	engagement	likes	comments	likes/followers	likes/mean	mean	Sunday	Monday	Tuesday	
0	<null>	<null>	400100000.0	7800000.0	2240988.0	18459.0	0.005601069732566858	0.33936452419598084	6603483.393879567	False	False	True	False
1	0.0	1.0	400100000.0	7800000.0	7886085.0	50874.0	0.019710284928767807	1.194231064063735	6603483.393879567	False	False	True	False
2	<null>	<null>	400100000.0	7800000.0	19643288.0	197619.0	0.04909594601349663	2.974685757248419	6603483.393879567	False	False	False	False
3	0.0	0.0	400100000.0	7800000.0	8363211.0	55938.0	0.020902801799550112	1.266484747694139	6603483.393879567	False	False	False	False
4	0.0	1.0	400100000.0	7800000.0	11434276.0	142142.0	0.028578545363659086	1.731552169964999	6603483.393879567	False	False	False	False
5	<null>	<null>	400100000.0	7800000.0	18780508.0	205000.0	0.04693953511622095	2.844030472978352	6603483.393879567	False	False	False	False
6	<null>	<null>	400100000.0	7800000.0	12838882.0	115864.0	0.03208918270432392	1.9442589969863038	6603483.393879567	False	False	False	False
7	0.0	0.0	400100000.0	7800000.0	16188034.0	606783.0	0.040459970007498125	2.4514385869439552	6603483.393879567	False	True	False	False
8	0.0	1.0	400100000.0	7800000.0	9332713.0	70339.0	0.023325951012246937	1.4133015021511248	6603483.393879567	True	False	False	False
9	0.0	0.0	400100000.0	7800000.0	10255392.0	106782.0	0.025632071982004498	1.553027605022101	6603483.393879567	False	False	False	False
10	0.0	2.0	400100000.0	7800000.0	8035532.0	56103.0	0.020083809047738065	1.2168626042806021	6603483.393879567	False	False	False	True
11	<null>	<null>	400100000.0	7800000.0	7873294.0	115298.0	0.019678315421144715	1.1922940560882391	6603483.393879567	False	False	False	False
12	2.0	2.0	400100000.0	7800000.0	7953689.0	37689.0	0.019879252686828292	1.204468691080812	6603483.393879567	False	False	False	False
13	0.0	0.0	400100000.0	7800000.0	6627776.0	28665.0	0.01656529867533117	1.0036787562974638	6603483.393879567	False	False	False	False
14	0.0	1.0	400100000.0	7800000.0	6463310.0	28936.0	0.01615423644088978	0.9787728104216198	6603483.393879567	False	False	False	True
15	1.0	1.0	400100000.0	7800000.0	5294226.0	38217.0	0.013232256935766058	0.8017323106933151	6603483.393879567	False	False	True	False
16	0.0	0.0	400100000.0	7800000.0	11170867.0	73474.0	0.027920187453136715	1.6916627685251253	6603483.393879567	True	False	False	False
17	0.0	1.0	400100000.0	7800000.0	2180145.0	20117.0	0.005449000249937516	0.3301507507417472	6603483.393879567	False	False	False	False
18	<null>	<null>	400100000.0	7800000.0	10752159.0	67939.0	0.026873679080229942	1.6282556279259564	6603483.393879567	False	False	False	True
19	0.0	4.0	400100000.0	7800000.0	9015981.0	40226.0	0.02253431892026993	1.3653371201563789	6603483.393879567	False	False	False	True
20	0.0	0.0	400100000.0	7800000.0	9017854.0	57651.0	0.022539000249937515	1.3656207583346374	6603483.393879567	False	False	False	True
21	<null>	<null>	400100000.0	7800000.0	8679568.0	54363.0	0.02169349662584354	1.3143923414791427	6603483.393879567	False	True	False	False
22	0.0	1.0	400100000.0	7800000.0	6951233.0	48526.0	0.017373739065233693	1.0526615401869177	6603483.393879567	False	False	False	False
23	<null>	<null>	400100000.0	7800000.0	4580145.0	31737.0	0.011447500624843789	0.6935952930910834	6603483.393879567	False	False	False	False

# Analiza datelor

## Matrice de confuzie

Importanța legăturii dintre caracteristici, raportate la numărul de like-uri pe media lor.

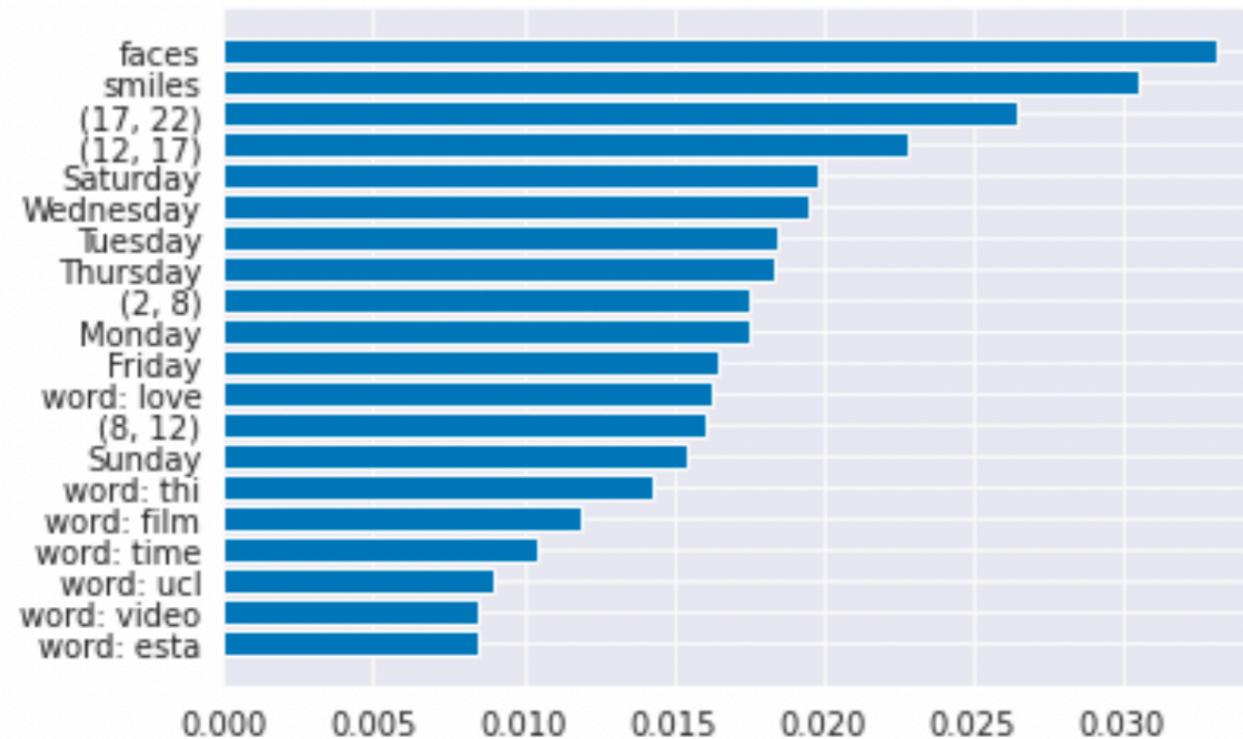


# Analiza datelor

## Importanța caracteristicilor

Cele mai importante caracteristici ce impactează numărul de like-uri primite în mod direct sunt:

- Numărul de fețe și zâmbete din imagini.
- Folosirea cuvântului 'love' sau alte forme ale acestuia, ex: 'loves', 'loving' etc.
- Orele cele mai bune de postat sunt între 17:00 și 22:00.
- Ziua cea mai bună în care să postăm este Sâmbătă.



Decision Tree Regression



Random Forest Regression

# Predictia postariilor

## Definirea modelului

- Vom folosi un model **Secvențial** cu 5 layere de  $215 * 1028$ ,  $1028 * 256$ ,  $256 * 128$ ,  $128 * 32$  și  $32 * 1$ . Între fiecare layer am adăugat câte o funcție de activare **ReLU**, iar la final am adăugat un dropout de 0.3 pentru a evita overfitting-ul.
- De asemenea, am folosit **L1Loss** și funcția de optimizare **Adam** cu *learning\_rate* = 0.0001.
- Dataset-ul va fi împărțit în 10% validation, 10% test și 80% train. Pentru fiecare dataframe, am creat batch-uri de câte 64.

# Predictia postariilor

## Evaluarea modelului

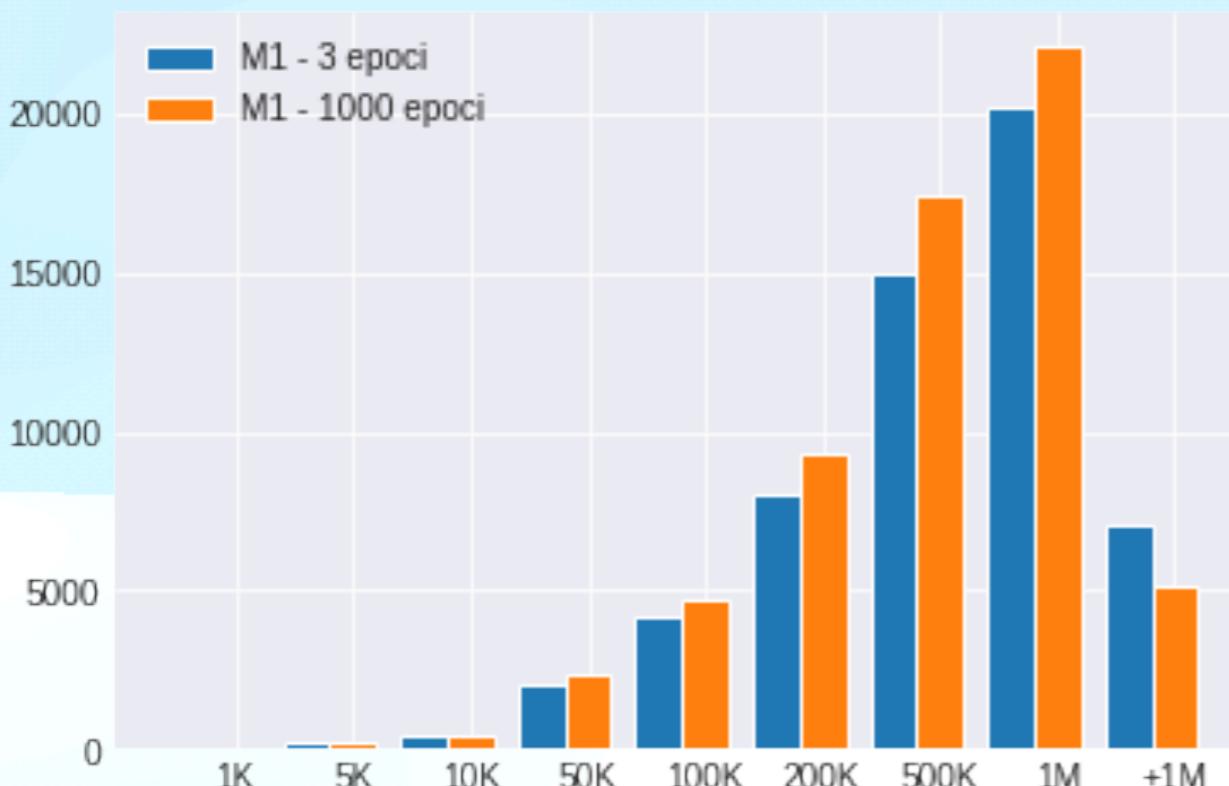
Pentru a putea evalua modelul nostru vom folosi următoarele metriki:

- $MSE = \frac{1}{n} * \sum_{i=1}^n (y_{true} - y_{predicted})^2$
- $MAE = \frac{1}{n} * \sum_{i=1}^n abs(y_{true} - y_{predicted})$ .
- Scorul  $R^2$

# Predictia postarilor

## Antrenarea modelului

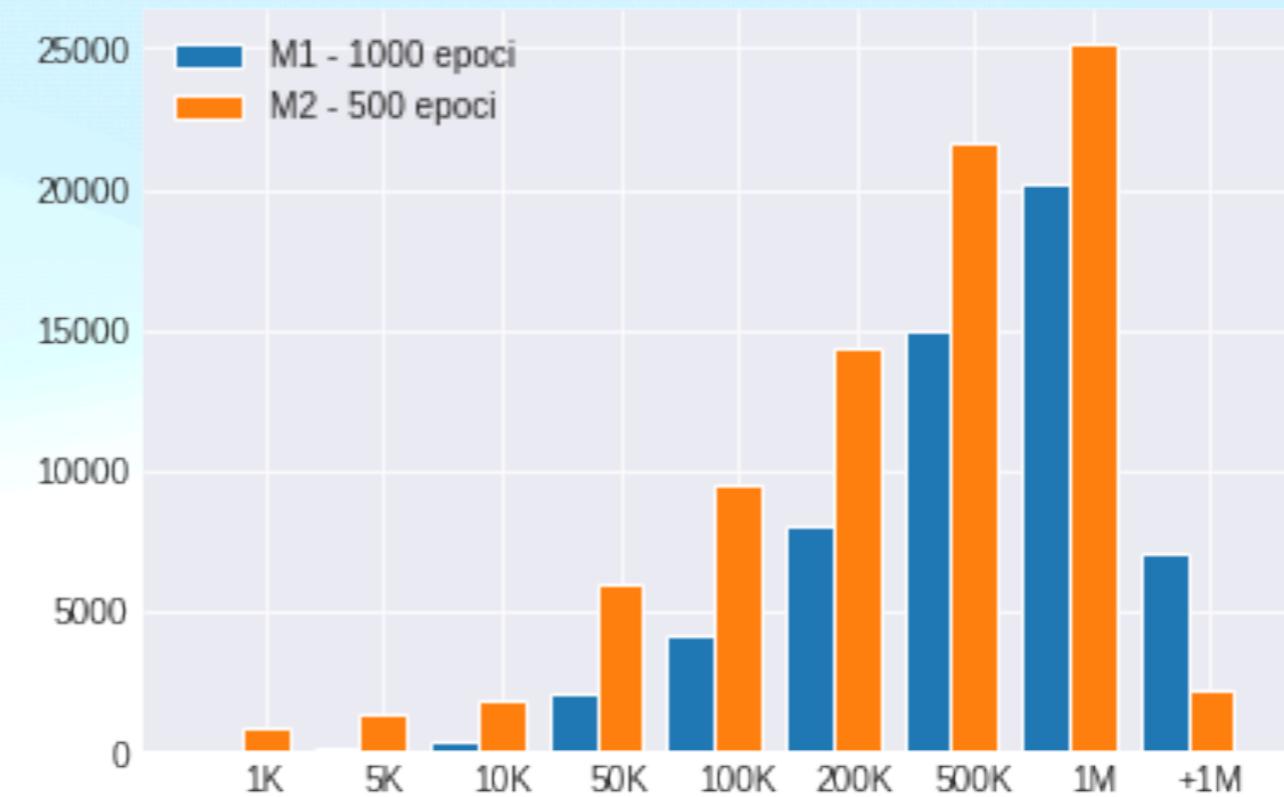
3 epochi vs 1000 epochi



Folosind likes/followers ca și metrică

MSE: 0.0011458757148844695  
Mean absolute error: 0.017063193946872352  
R2 score: -16.014154436814696

Model 1 vs Model 2



Folosind likes/mean ca și metrică

MSE: 0.776016050386604  
Mean absolute error: 0.5172142815730558  
R2 score: -2.0881367596652693

# Aplicația mobilă



## Tehnologii folosite

Pentru a realiza aplicația mobilă vom folosi următoarele tehnologii:

- **Flask** pentru Backend
- **MySQL** pentru baza de date
- **Flutter** pentru Frontend - Aplicația mobilă



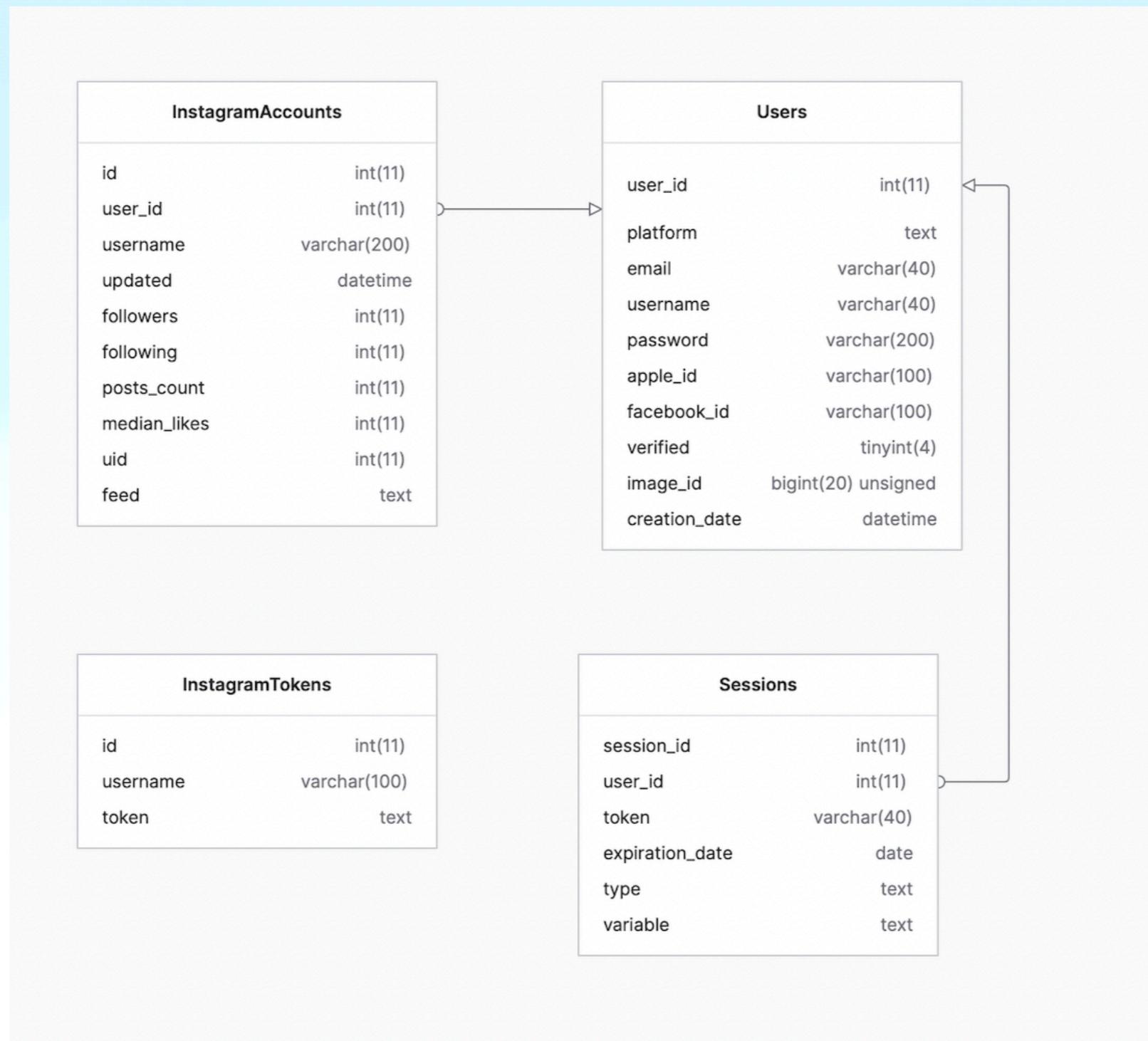
Flask



# Aplicația mobilă



## Baza de date

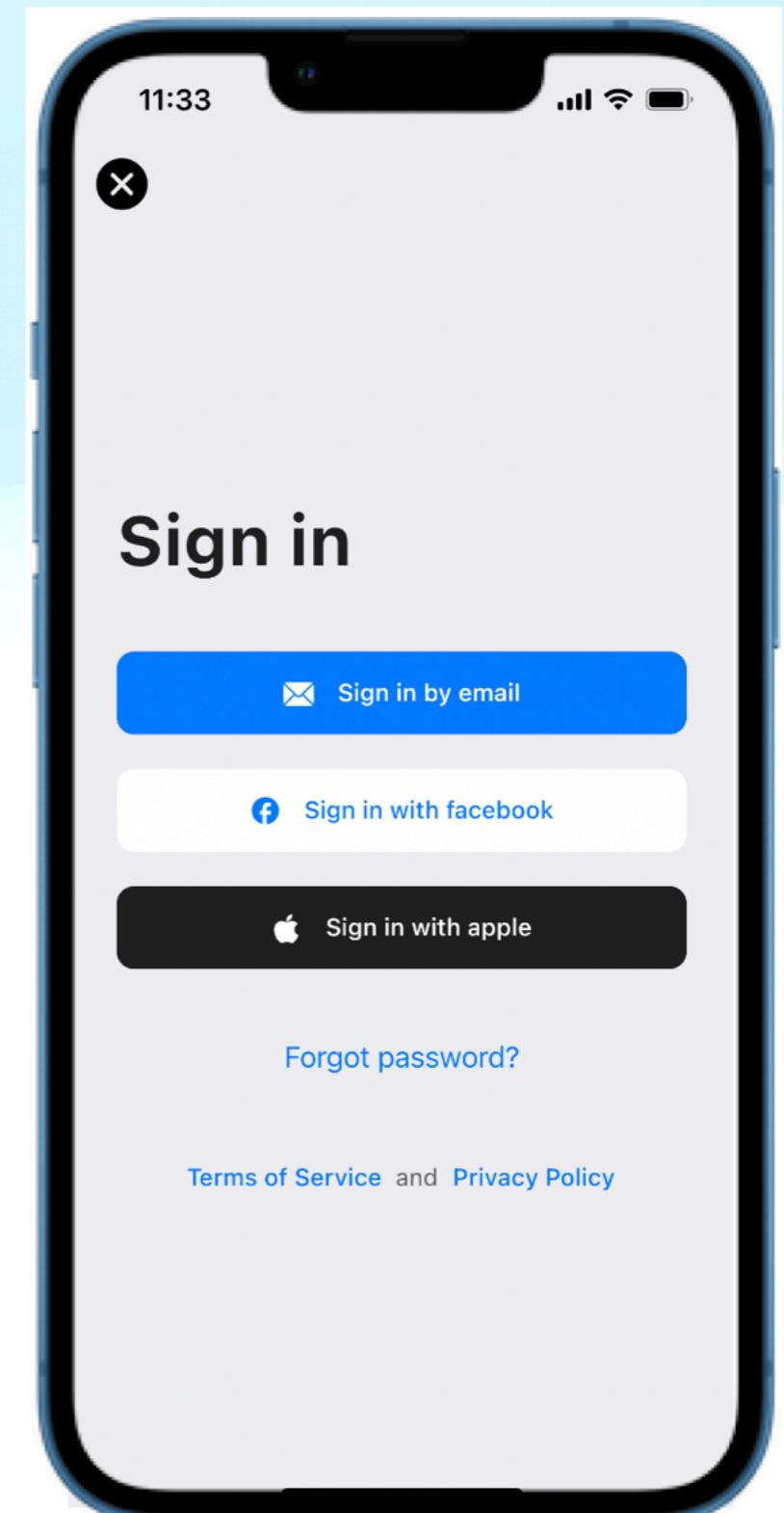


# Aplicația mobilă



## Autentificarea

- Autentificarea cu Apple folosind librăria **sign\_in\_with\_apple**.
- Autentificarea cu Facebook folosind librăria **flutter\_facebook\_auth**.
- Autentificarea prin **Email**.

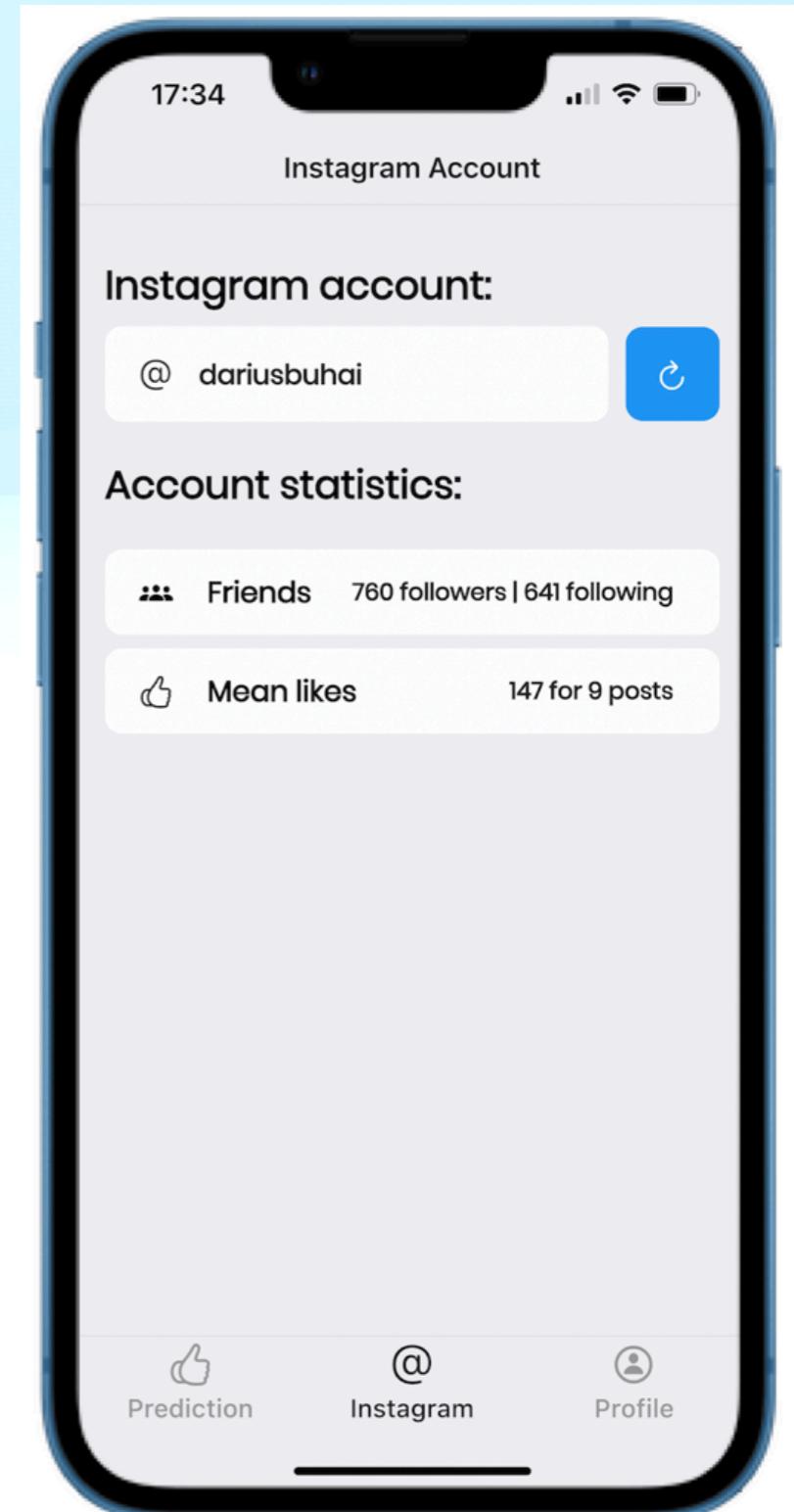


# Aplicația mobilă

## Contul de Instagram



- Pentru a prelua media like-urilor precedente și numărul de urmăritori, vom conecta contul de instagram al utilizatorului la contul său din aplicație.
- Desigur, pentru a avea access la statisticile contului, acesta trebuie să fie public și să conțină cel puțin 1 postare.

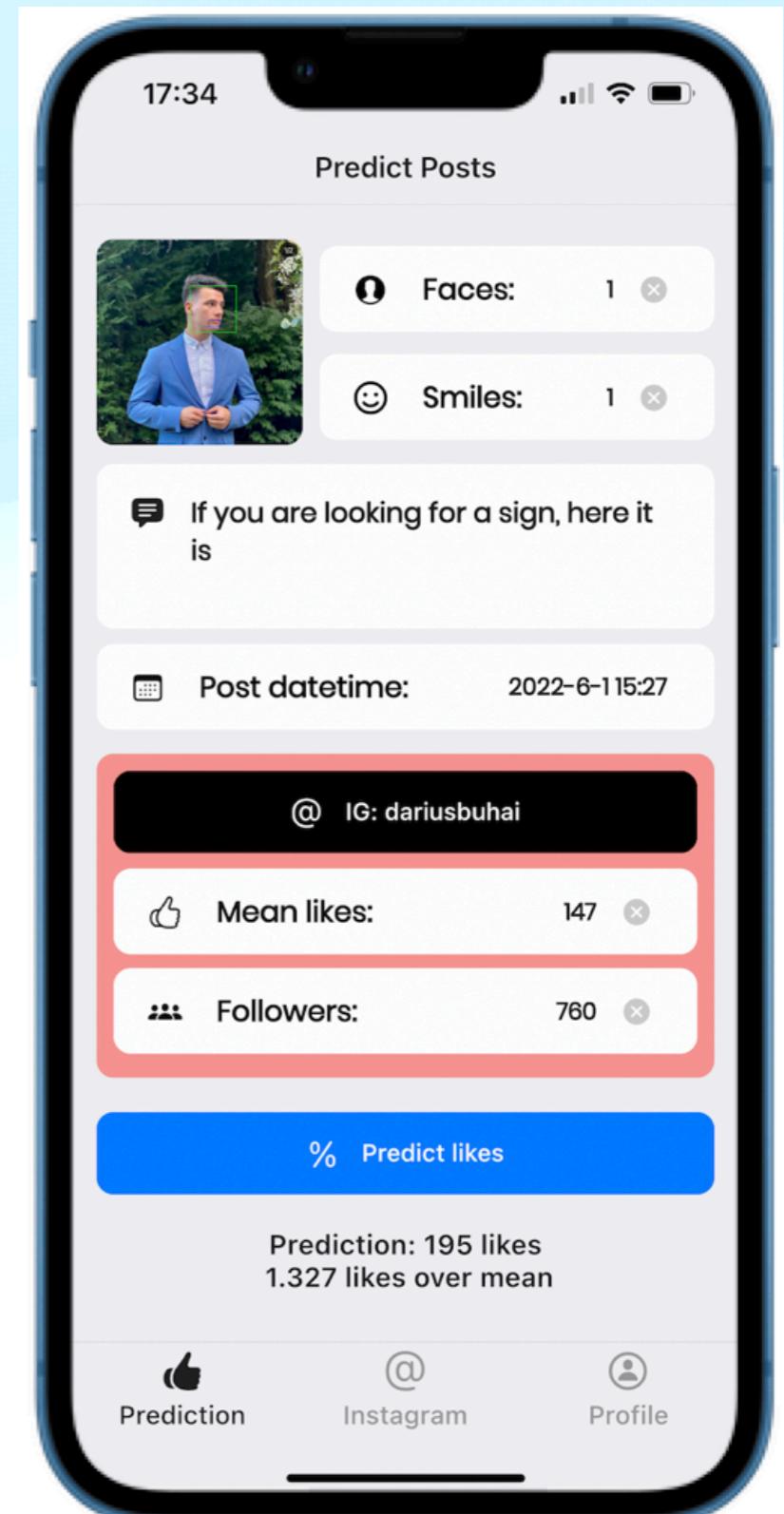


# Aplicația mobilă

## Predictia postărilor



- Utilizatorul încarcă imaginea.
- Sunt complete datele referitoare la postare (descriere, dată și ora).
- Se preiau automat detalii referitoare la contul de instagram al utilizatorului.
- Se afișează predictia.



# Concluzie

## Îmbunătățiri

- Mai multe caracteristici de antrenare a modelului.
- Un set de date mai mare și mai diversificat.
- Antrenarea modelului trebuie să fie specifică pentru fiecare cont în parte și actualizat constant.

# Demo

