Uputstvo za testere

Projekat igrANNonica Tim RevolutionN

Sadržaj

Korisni podaci	2
Opis slučajeva korišćenja	2
Početni prikaz aplikacije	2
Logovanje na sistem	2
Registracija na sistem	5
Treniranje modela bez potrebe za prijavom	6
Izbor dataseta	7
Podešavanje dataseta	8
Obuka modela	10
Pregled uspešnosti modela	14
Treniranje modela kada je korisnik prijavljen	15
Kreiranje datasetova	15
Pregled kreiranih datasetova	17
Korišćenje sačuvanih datasetova	18
Kreiranje modela	19
Upoređivanje modela	20
Predikcija	21
Pregled My Profile stranice	23
Izmena podataka	23
Brisanje korisničkog naloga	24
Odjava sa sistema	24
Learn More stranica	25

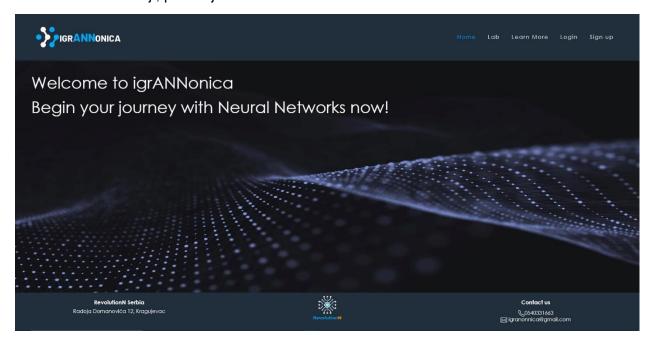
Korisni podaci

- IP adresa sajta: http://softeng.pmf.kg.ac.rs:10135/
- Nalozi koji se mogu koristiti bez potrebe za registracijom: username:user, password:user username:korisnik , password:korisnik

Opis slučajeva korišćenja

Početni prikaz aplikacije

Prilikom ulaska na sajt, prikazuje se Home stranica:

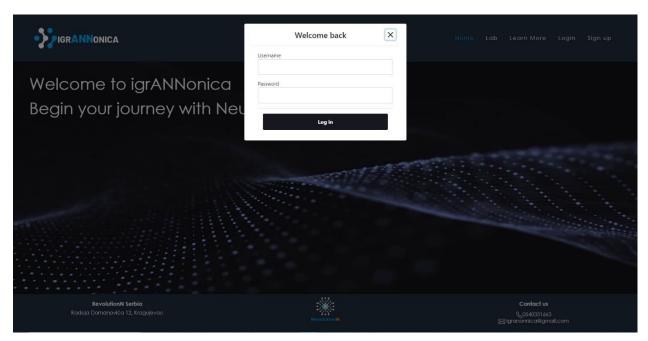


Korisnik može izabrati različite opcije iz menija.

- Home stranica na kojoj se trenutno nalazi
- Lab stranica na kojoj čak i neprijavljen korisnik može izvršiti treniranje modela nad nekim od ponuđenih datasetova
- Learn More stranica na kojoj se korisnik može informisati o veštačkim neuronskim mrežama, i takođe prođi kroz korake koji su potrebni za kreiranje jednog modela za predikciju
- Login ukoliko korisnik želi da se prijavi na sistem već postojećim nalogom
- Sign up ukoliko korisnik želi da se registruje i napravi nov nalog

Logovanje na sistem

Klikom na Login u navigacionom meniju, korisniku se prikazuje sledeći modal:



Ukolilko korisnik unese ispravne podatke, prikazaće mu se obaveštenje o uspešnoj prijavi na sistem, sa porukom Welcome, <username>.

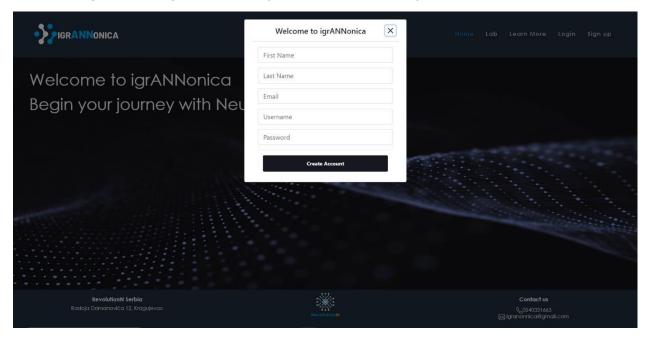


Ukoliko pak korisnik ne unese dobre podatke, prikazaće mu se poruka da prijava na sistem nije uspešna, jer username ili password nisu ispravni.

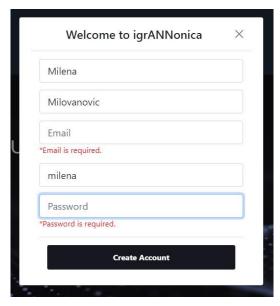


Registracija na sistem

Klikom na Sign up u navigacionom meniju, korisniku se prikazuje sledeći modal:



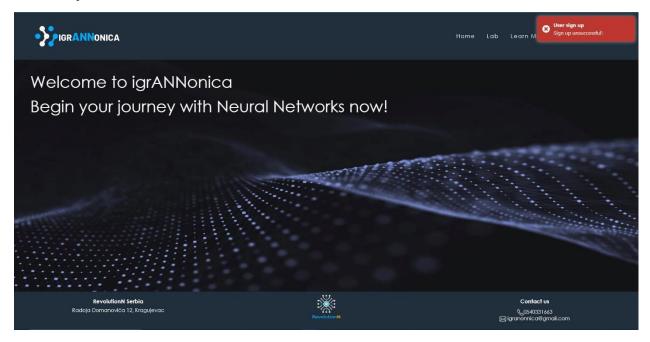
Ukoliko korisnik neko polje ostavi prazno, prikazaće mu se upozorenja da su ta polja obavezna, i ako želi da se registruje mora i njih da popuni.



Nakon uspešne registracije, korisnik automatski biva i ulogovan, kako bi se ubrzalo vreme prijave na sistem. Korisnik dobija sledeće obaveštenje:

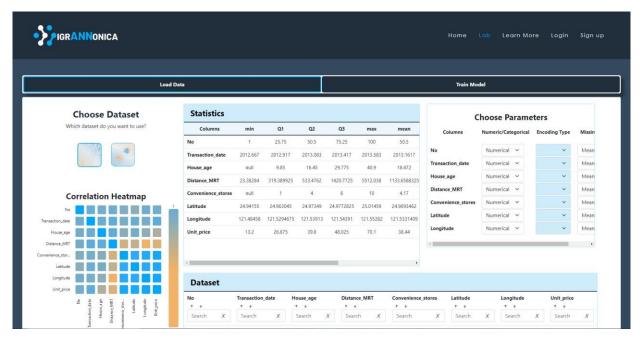


Ukoliko je registracija korisnika na sistem iz nekog razloga neuspešna, prikazuje se sledeće obaveštenje:



Treniranje modela bez potrebe za prijavom

Korisnik može izvršiti treniranje modela i u slučaju kada ne želi da kreira nalog na sajtu. Klikom na stranicu Lab u navigacionom meniju, otvara se sledeća stranica:



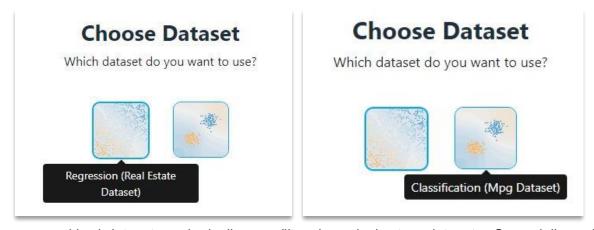
Korisniku je omogućeno da izabere jedan od dva ponuđena dataseta, podesi različite parametre za svaku od kolona tog dataseta i nad njim istrenira neku neuronsku mrežu. Na kraju treninga može da vidi grafike uspešnosti tog modela.

Međutim, ovakav korisnik ne može da vrši dodavanje svojih datasetova, predikciju istreniranim modelom, niti da čuva bilo kakve podatke.

Izbor dataseta

Ono što prvo biva prikazano korisniku je Load Data stranica. Na njoj korisnik može da bira između dva dataseta koji su ponuđeni. Na stranici je automatski učitan prvi dataset, tako da korisnik odmah može da pređe na konfiguraciju parametara, ili da odabere drugi ponuđen dataset.

Prvi dataset je predviđen za regresioni problem, a drugi za klasifikacioni. Prelaskom miša preko dataseta korisnik dobija informacije koji je dataset u pitanju i za koji tip problema je namenjen.



Klikom na neki od datasetova, korisniku se učitavaju podaci o tom datasetu. On nadalje može vršiti njegovo podešavanje.

Podešavanje dataseta

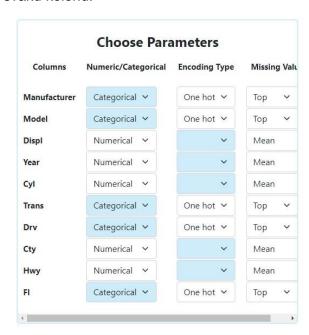
Korisnik za učitani dataset može vršiti podešavanje određenih parametara.

1. Statistika

Za svaku kolonu korisnik ima uvid u statističke podatke:

Columns	min	Q1	Q2	Q3	max	mean	std	count	freq	to
Manufacturer	null	null	null	null	null	null	null	100	37	do
Model	null	null	null	null	null	null	null	100	37	cara 2\
Displ	1.6	2.4	2.4	4.6	5.3	3.385	1.3885807739	100	null	n
Year	1999	1999	1999	1999	2008	2000.71	3.5484951996	100	null	n
Cyl	4	4	4	8	8	5.76	1.9955506063	100	null	n
Trans	null	null	null	null	null	null	null	100	44	auto
Drv	null	null	null	null	null	null	null	100	56	
Cty	11	13.25	18	18	28	15.59	3.253886565	100	null	n
Hwy	17	19.25	24	24	33	22.48	4.2981790836	100	null	n
								2		-

2. Izbor parametara za svaku kolonu:

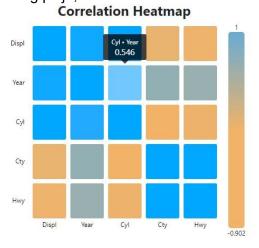


Za svaku kolonu korisnik može izabrati:

- Numeric/Categorical da li kolonu posmatra kao numeričku ili kategoričku. Podrazumevano će kolone koje se sastoje od stringova biti Categorical, a one koje se sastoje od brojeva biti Numerical. Međutim, korisnik može promeniti one kolone koje su Numerical u Categorical (na primer ako su vrednosti 0,1,2 – želi da ih posmatra kao kategorije).
- Encoding Type enkodiranje kolona može izabrati samo za vrednosti koje su podrazumevano kategoričke, jer su one Stringovi. Može izabrati One hot, Label i Binary. Podrazumevano će biti odabran One hot.
- Missing Values I za numeričke i za kategoričke se može birati način popunjavanja nedostajućih vrednosti. Numeričke kolone se mogu popunjavati sa Mean, Median, Minimum, Maximum i Delete(ne uzimati u obzir tu kolonu u predikciji). Kategoričke kolone se mogu popunjavati sa Top(vrednost koja se najčešće ponavlja) i Delete.

3. Korelaciona matrica u vidu heatmape

Za svaki par kolona korisnik može imati uvid u njihovu korelaciju odnosno zavisnost. Sa strane se nalazi legenda, koja pokazuje boje kojima su obojene različite zavisnosti. Prelaskom mišem preko nekog polja, vidi se tačna zavisnost te dve kolone.



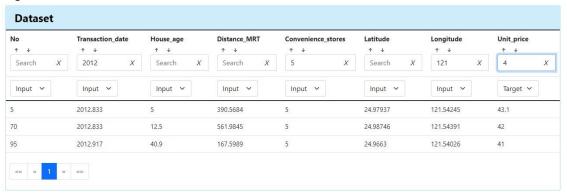
4. Tabela samog dataseta

No ↑ ↓ Search X	Transaction_date ↑ ↓ Search X	House_age ↑ ↓ Search X	Distance_MRT ↑ ↓ Search X	Convenience_stores ↑ ↓ Search X	Latitude ↑ ↓ Search X	Longitude ↑ ↓ Search X	Unit_price ↑ ↓ Search X
Input ~	Input ~	Input ~	Input ~	Input ~	Input ~	Input ~	Target V
1	2012.917	32	84.87882	10	24.98298	121.54024	37.9
2	2012.917	19.5	306.5947	9	24.98034	121.53951	42.2
3	2013.583	13.3	561.9845	5	24.98746	121.54391	47.3
4	2013.5	13.3	561.9845	5	24.98746	121.54391	54.8
5	2012.833	5	390.5684	5	24.97937	121.54245	43.1
6	2012.667	7.1	2175.03	3	24.96305	121.51254	32.1
«« « 1 2	3 4 5 6 7	8 9 10 11	12 13 14 15	16 17 » »»			

Na kraju, korisnik može videti i sam dataset. Klikom na strelice ispod naziva kolone da sortira dataset rastuće ili opadajuće po vrednostima te kolone.

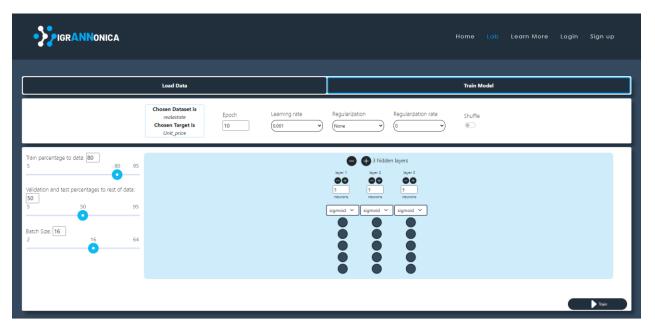
Svaka kolona nudi Search polje, i korisnik može zadati zasebno parametre za pretragu svake od kolona. U tabeli će se prikazati samo vrste koje zadovoljavaju sve unesene parametre za pretragu.

Takođe, za svaku kolonu postoji padajući meni koji predstavlja izbor da li je kolona Input, odnosno uključena u trening ili Target odnosno ono što model treba da prediktuje. Takođe, može izabrati i vrednost None, čime će isključiti kolonu iz dataseta. Ove vrednosti će podrazumevano biti takve da su sve kolone sem poslednje Inputi, a poslednja kolona Target.



Obuka modela

Klikom na dugme Train Model korisnik prelazi na deo sa podešavanjem hiperparametara samog modela.

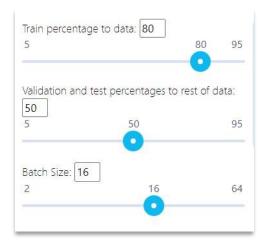


Svi hiperparametri imaju podrazumevane vrednosti, kako bi korisnik što lakše mogao da pokrene obuku uz samo par podešavanja.

Chosen Dataset is realestate	Epoch	Learning rate	Regularization	Regularization rate	Shuffl
Chosen Target is	10	0.001	None	0 •	
Unit_price					

Korisnik može da podesi broj epoha treninga, learning rate, tip regularizacije, regularization rate i shuffle(da li će podaci biti u redosledu koji je podrazumevan ili će se redosled vrsta izmešati).

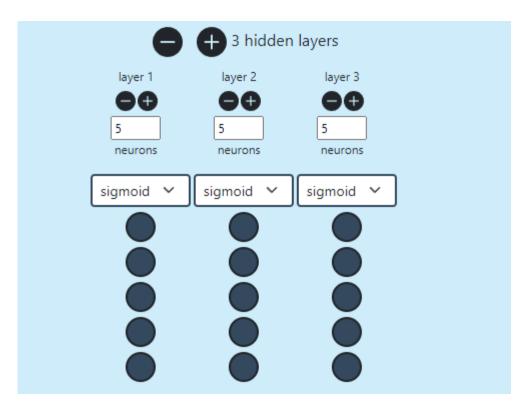
Ono o čemu treba voditi računa je Problem type, koji predstavlja tip problema kojim se bavimo. Može biti Regression ili Classification. Izbor ovoga zavisi od Targeta dataseta koji smo izabrali u prethodnom koraku.



Sledeće što korisnik može da bira je:

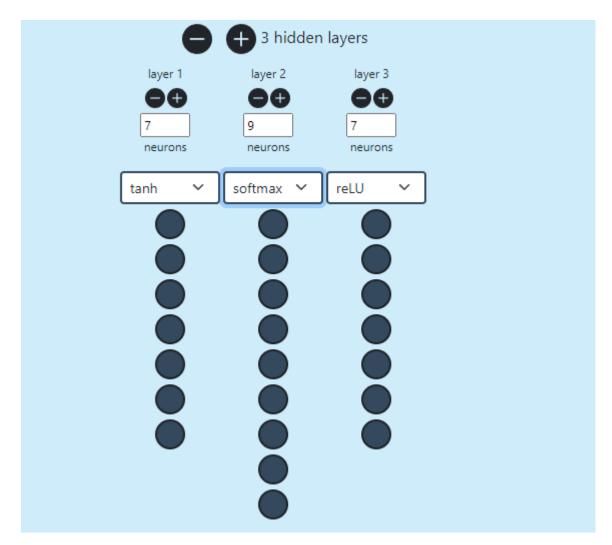
- Train percentage to data koji procenat podataka će se uzeti za obuku modela
- Validation and test percentages to rest of data odnos validacionog i testnog skupa, koji će se kreirati od ostatka koji nije uzet za trening
- Batch size veličina trening uzoraka koji će biti korišćeni u jednoj iteraciji treninga

Poslednja stvar koju korisnik treba da podesi je izgled same mreže odnosno broj slojeva i neurona unutar svakog sloja. Minimalni broj slojeva je 1, a maksimalni 7. Minimalni broj neurona po sloju je 1. Podrazumevani izgled je sledeći:

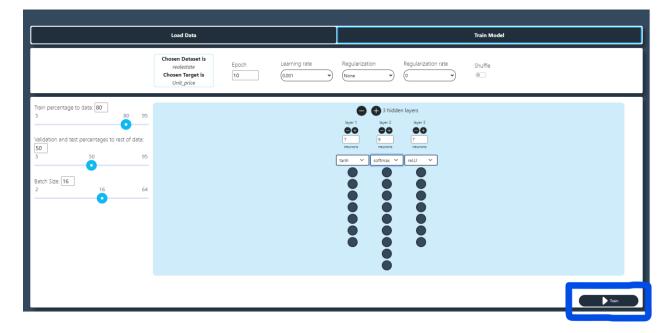


Klikom na plus/minus pored broja skrivenih slojeva korisnik može dodavati još slojeva ili brisati poslednji. Takođe klikom na plus/minus za svaki sloj može dodati ili brisati neurone. Za svaki sloj zasebno može odabrati aktivacionu funkciju (Sigmoid, Linear, Tanh, Softmax ili reLU). Podrazumevana je uvek Sigmoid.

Dakle, izgled mreže posle par izmena može biti ovakav:



Na kraju, klikom na dugme train u donjem desnom uglu, korisnik pokreće proces obuke:



Tokom trajanja obuke, korisniku se prikazuje sledeći loader, koji označava da je trening u toku.



Pregled uspešnosti modela

Kada model završi obuku, korisniku se prikazuju grafici različiti u zavisnosti od toga da li je problem klasifikacioni ili regresioni.

U slučaju klasifikacionog problema, grafici će biti sledeći:

Sa leve strane korisnik može čekirati koje grafike želi/ne želi da vidi. Prva dva grafika se prikazuju uvek, a to su grafik evaluacije metrika, i grafik na kom se vide stvarne vrednosti i one koje je prediktovao istrenirani model.

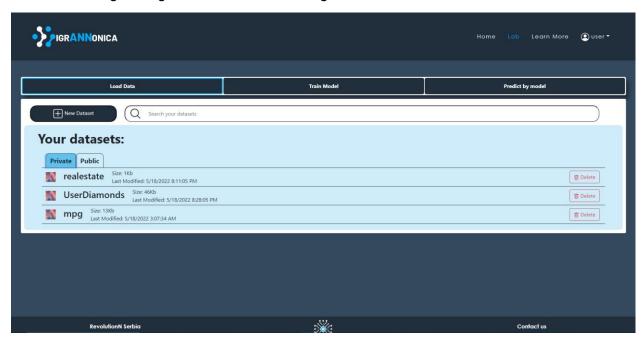
Na sledećoj slici prikazan je izgled grafika kada korisnik odabere šta želi da vidi.



Treniranje modela kada je korisnik prijavljen

Korisnik koji je prijavljen ima par funkcionalnosti više od korisnika koji nije prijavljen. One se tiču najviše pamćenja rezultata i njihovog ponovnog korišćenja.

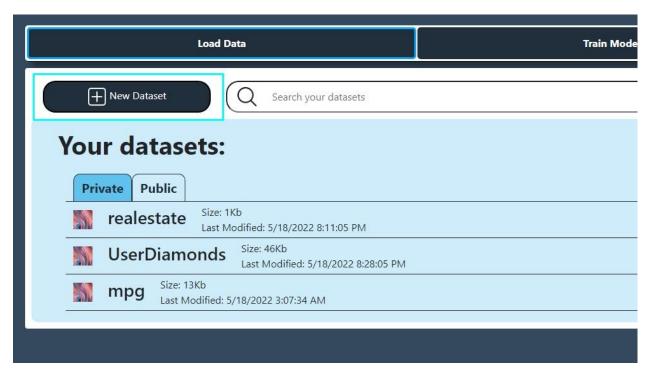
Lab stranica ulogovanog korisnika ima sledeći izgled:

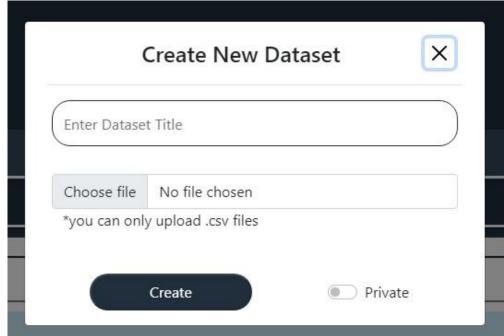


Korisnik ima mogućnost prikaza 3 stranice – Load Data, Train Model (koje ima i neulogovan) i Predict by Model gde može koristiti obučene modele za predikciju.

Kreiranje datasetova

Prva funkcionalnost koju ulogovani korisnik ima je kreiranje datasetova. Klikom na dugme New Dataset mu se otvara modal za kreiranje novog dataseta.

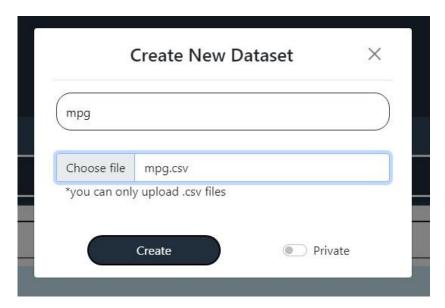




Korisnik mora uneti naziv dataseta i csv fajl koji predstavlja dataset. Dok ne unese oba, nije mu dozvoljeno da klikne na Create.

Takođe, korisnik ima i Private/Public toggle, kojim može da izabere da li će uneti dataset da čuva kao privatni, ili želi da on bude dostupan i drugim korisnicima.

Ukoliko korisnik nije uneo naziv dataseta, a uploaduje fajl, automatski će mu se polje za naziv popuniti nazivom fajla koji je dodao.



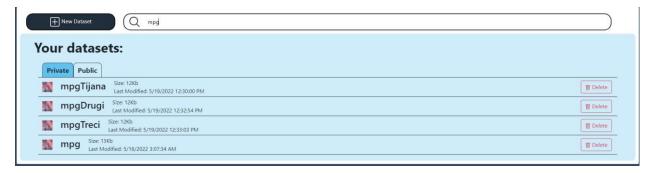
Pregled kreiranih datasetova

Korisnik može da koristi datasetove koje je kreirao on, ili one koji su javni. Ti datasetovi prikazani su u vidu tabele, i razdvojeni u Private i Public delove.



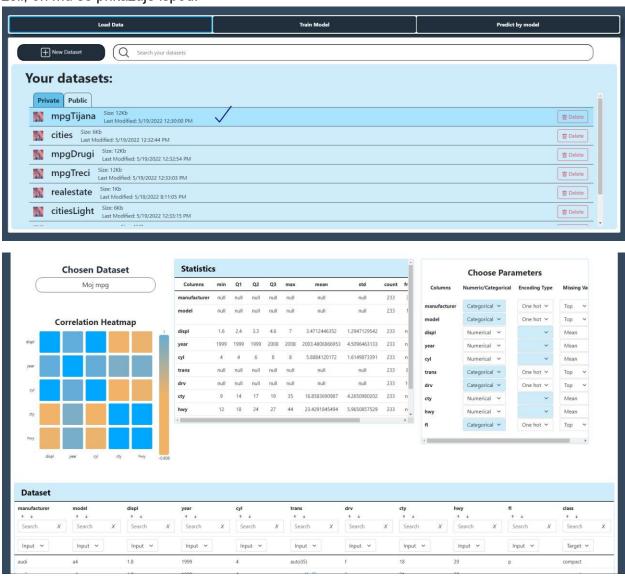
Za Private datasetove korisnik ima opciju Delete, dok za Public nema. Klikom na dugme Delete briše se dataset, kao i svi modeli koji su nad tim datasetom istrenirani.

I za privatne i za javne datasetove korisnik ima mogućnost pretrage, putem searchbar-a.



Korišćenje sačuvanih datasetova

Korisnik može koristiti bilo koji Private ili Public dataset za obuku modela. Klikom na dataset koji želi, on mu se prikazuje ispod.



Sada, korisnik ima sve iste mogućnosti koje je imao neulogovan korisnik. Te funkcionalnosti opisane su <u>ovom</u> odeljku.

Kreiranje modela

Kada korisnik podesi sve što je želeo nad određenim datasetom, klikom na Train Model dugme prelazi na sledeću stranicu na kojoj vrši podešavanje hiperparametara i strukture modela koji želi da istrenira.

Takođe i ovde korisnik ima iste mogućnosti koje je imao neulogovan korisnik u odeljku <u>Obuka modela</u> i <u>Pregled uspešnosti modela</u>.

Međutim, jedina razlika je u tome što kada se obuka završi, korisnik može sačuvati ovaj model ukoliko je zadovoljan njegovim rezultatima. Taj model on kasnije može koristiti za predikciju. Korisnik treba da popuni polje za naziv dataseta, i može izabrati da li će taj model biti privatni ili javni. Klikom na dugme save korisnik dobija obaveštenje da li je uspešno sačuvan model ili ne.

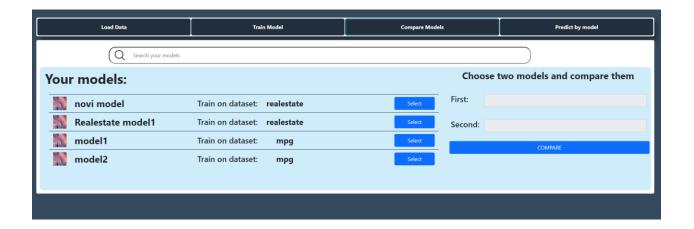


Ukoliko je model sačuvan uspešno, izlazi sledeće obaveštenje:



Upoređivanje modela

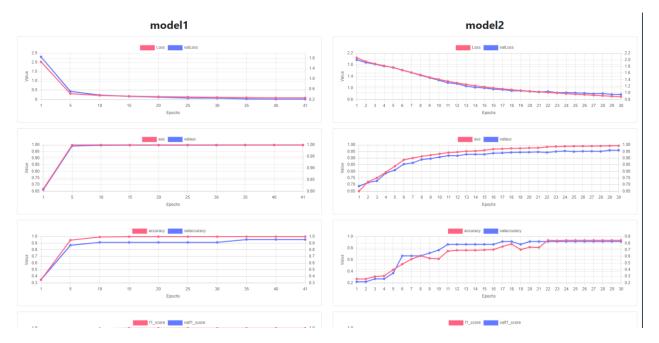
Takođe, ulogovani korisnik ima mogućnost i da uporedi dva sačuvana modela koje je kreirao.



Korisnik klikom na dugme "Select" čekira željeni model, koji će se upisati u textboxu i takođe će se dugme kliknutog modela onemogućiti.

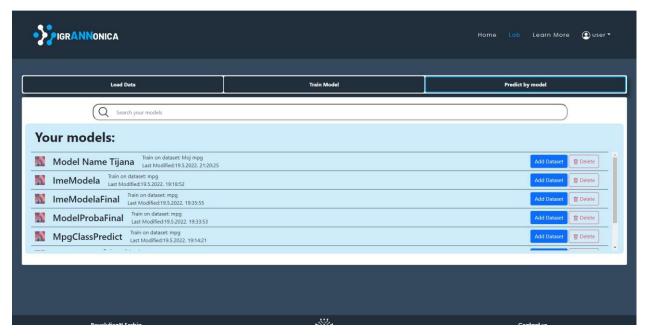


Konačno, klikom na dugme "Compare", prikazaće se grafici odnosno vizuelizacija obuke za oba modela.



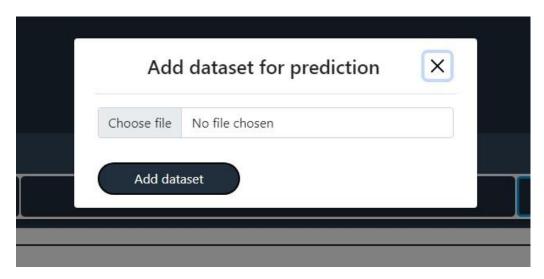
Predikcija

Još jedna funkcionalnost koju ulogovan korisnik ima je predikcija. Klikom na dugme Predict by model, korisnik odlazi na sledeću stranicu:



Korisnik može vršiti pretragu svojih modela putem Search bar-a. Takođe ima uvid u to nad kojim datasetom je istreniran model. Svaki model može i obrisati klikom na dugme Delete.

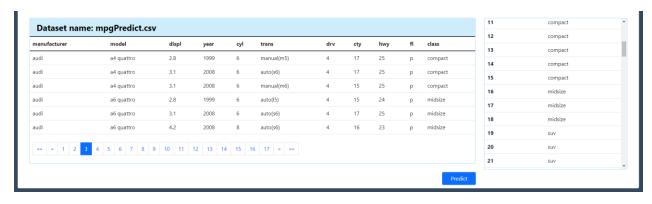
Klikom na dugme Add Dataset, korisniku se otvara sledeći modal:



Nakon što korisnik unese dataset za predikciju, on se prikazuje u vidu tabele:

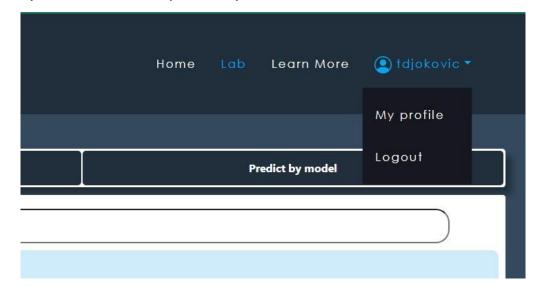


Klikom na dugme Predict se pokreće predikcija putem odabranog modela nad postavljenim datasetom. Sa desne strane se za svaki red prikazuje prediktovana vrednost.

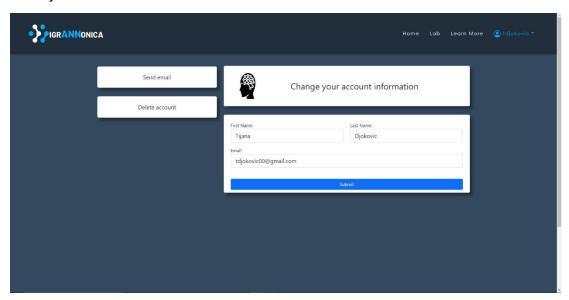


Pregled My Profile stranice

Klikom na username u navigacionom meniju, korisniku se otvara dropdown koji mu daje izbor pregleda My Profile stranice, i odjave sa sajta.



Klikom na My Profile otvara se sledeća stranica:



Izmena podataka

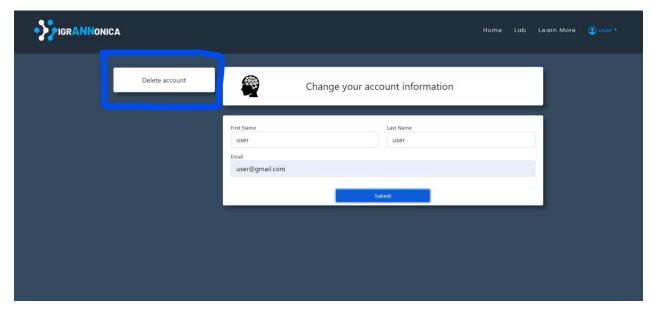
Korisnik može da izmeni Ime, Prezime ili Mejl adresu koju je uneo prilikom registracije.

Change your account information					
First Name Tijana	Last Name Djokovic				
Email tijana@gmail.	com				
	Submit				

Nakon izmene, klikom na dugme Submit izmene bivaju sačuvane.

Brisanje korisničkog naloga

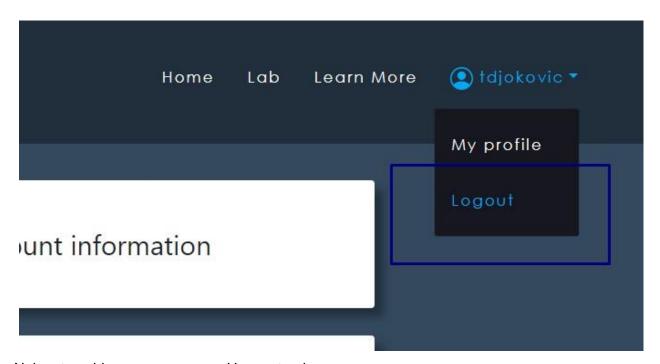
Korisnik ima moćusnost da svoj nalog obriše klikom na Delete account dugme sa leve strane.



Nakon toga automatski biva izlogovan i preusmeren na Home stranicu.

Odjava sa sistema

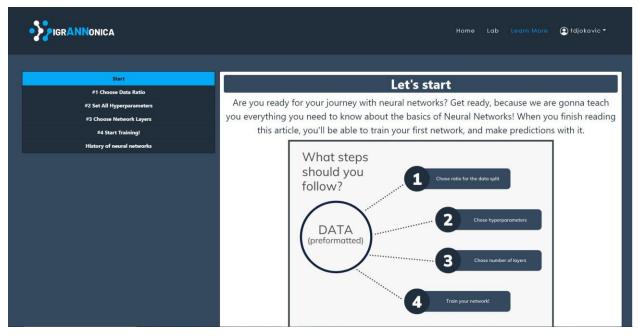
Klikom na Logout dugme u padajućem meniju navigacionog menija, korisnik može da se odjavi.



Nakon toga biva preusmeren na Home stranicu.

Learn More stranica

Learn More je informativna stranica, gde korisnici koji nemaju neko predznanje o Veštačkim neuronskim mrežama mogu da nauče osnovne korake potrebne za kreiranje nekog modela za predikciju.



Sa leve strane se nalazi meni koji pokazuje korake koji su obavezni za trening jednog modela. Klikom na bilo koju stavku, prikaz prelazi na taj deo stranice.

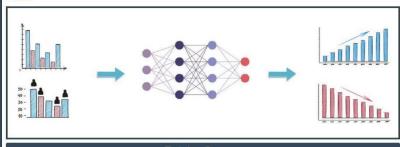
Start
#1 Choose Data Ratio
#2 Set All Hyperparameters
#3 Choose Network Layers
#4 Start Training!
History of neural networks

Neural Network Layers

Now, you need to choose number of layers of your model, and number of neurons that each one of them will contain.

These are the layers that come between the input and output layers.

Input layer is the one accepting plain data as input, and the output layer is the one capturing inputs from hidden layer before itself, performs the calculations via its neurons and then the output is computed. The presence of more hidden layers normally improves accuracy to a degree that can change depending on the problem. Each one of hidden layers that you add can have different number of neurons, and separate activation functions. This gives you flexibility for more specified construction of the network.



Training Process

Now comes the *fun part* - training and testing the model you've just made!

When the model is done training on the data you've provided, you can see multiple metrics that evaluate the success of the data