

# Prepoznavanje matematičkih izraza

Autor: Jelena Popov RA 29/2016

Asistent: Stefan Anđelić

## Motivacija

Sve je veća potreba za digitalnom predstavom matematičkih izraza radi njihovog kasnijeg izračunavanja pomoću računara. Prepoznavanje rukom pisanih matematičkih izraza omogućava prenošenje izraza sa papira u računar kako bi se dalje iskoristili u pisanju naučnih radova i druge literature.

## Skup podataka

Kao obučavajući skup korišćene su slike pojedinačnih matematičkih simbola i brojeva. Primeri nekih simbola su dati ispod.

+ α 2 < φ ÷

Model je testiran na četiri grupe slika: slike koje sadrže jedan matematički izraz, slike koje sadrže više izraza, slike koje sadrže jedan zarotiran izraz i slike koje sadrže više zarotiranih izraza. Primer jednog izraza iz test skupa je dat ispod.

$$151 \pm 143 \div 97$$

## Obučavanje modela

Za obučavanje modela koji će prepoznavati izdvojene matematičke izraze koristila sam dva pristupa:

1. CNN sa 9 slojeva - 8 skrivenih i jednim izlaznim slojem
2. LSTM sa 6 slojeva - 5 skrivenih i jednim izlaznim slojem.

## Pretprocesiranje podataka

### Slike sa jednim matematičkim izrazom

1. Prvi korak je učitavanje slike

$$151 \pm 143 \div 97$$

2. Nakon toga je primenjen threshold i pronađene su konture

$$151 \pm 143 \div 97$$

3. Grupišu se konture koje u obučavajućem skupu predstavljaju jednu sliku

$$151 \pm 143 \div 97$$

4. Od svakog bounding box-a koji obuhvata jedan simbol pravi se slika veličine 28x28 koja se šalje mreži za predikciju

$$151 \pm 143 \div 97$$

### Slike sa više matematičkih izraza

1. Prvi korak je učitavanje slike

$$\begin{aligned} (2-1-1) \\ 10x=3 \\ b \neq 0 \end{aligned}$$

2. Nakon toga je primenjen threshold pa dilacija kako bi se konture simbola jednog matematičkog izraza spojile u jednu

$$\begin{aligned} (2-1-1) \\ 10x=3 \\ b \neq 0 \end{aligned}$$

3. Zatim su primenjeni koraci kao za slike sa jednim matematičkim izrazom nad svim uočenim izrazima pojedinačno

$$\begin{aligned} (2-1-1) \\ 10x=3 \\ b \neq 0 \end{aligned}$$

### Slike sa jednim zarotiranim matematičkim izrazom

1. Prvi korak je učitavanje slike

$$1 - e^1$$

2. Nakon toga je uočen matematički izraz i određen njegov ugao zarotiranosti koji je upotrebljen da se matematički izraz ispravi

$$1 - e^1$$

3. Zatim su primenjeni koraci kao za slike sa jednim matematičkim izrazom nad svim uočenim izrazima pojedinačno

### Slike sa više zarotiranih matematičkih izraza

1. Prvi korak je učitavanje slike

$$\begin{aligned} 3n^2 + 2n \\ x = \frac{1}{y} \\ 2e^1 - 3e - 2 - 3i = 0 \end{aligned}$$

2. Nakon toga je primenjen isti postupak kao za slike sa više matematičkih izraza da bi se izdvojili pojedinačni izrazi.
3. Nad svakim pojedinačnim matematičkim izrazom je primenjen isti postupak kao za slike sa jednim zarotiranim izrazom kako bi se ispravio.
4. Zatim su primenjeni koraci kao za slike sa jednim matematičkim izrazom nad svim uočenim izrazima pojedinačno

## Formiranje matematičkog izraza

Nakon što mreža vrati predikcije za pojedinačne matematičke simbole potrebno je od tih predikcija formirati matematički izraz. Ono što daje težinu ovom koraku su izrazi sa:

1. Razlomcima, u kojima treba odrediti brojilac i imenilac
2. Korenom, u kojima treba odrediti koji šta je pod korenom
3. Indeksima i eksponentima, u kojima treba odrediti šta je u indeksu/eksponentu

Primer formiranja matematičkog izraza:

1. Učitavanje i pretprocesiranje slike koja je prikazana ispod

$$\frac{2}{3} n^3$$

2. Predikcija modela za pojedinačne simbole je [2, 3, n, 3]
3. Formirani matematički izraz je  $\frac{2}{3} n^3$

## Određivanje tačnosti

Radi utvrđivanja tačnosti modela korišćene su:

1. Scikit-learn accuracy\_score za proveru tačno prediktovanih matematičkih izraza
2. Prosečan FuzzyWuzzy ratio za utvrđivanje procenta poklapanja prediktovanih i tačnih matematičkih izraza

## Rezultati

Oba primenjena pristupa za prepoznavanje matematičkih simbola (CNN i LSTM) daju približne rezultate za sve test primere. Ispod su date prosečne tačnosti za sve četiri grupe slika.

1. Za slike koje sadrže jedan matematički izraz:
  - accuracy\_score: 60
  - prosečan FuzzyWuzzy ratio: ~91
2. Za slike koje sadrže više matematičkih izraza:
  - accuracy\_score: ~60
  - prosečan FuzzyWuzzy ratio: ~91
3. Za slike koje sadrže jedan zarotirani matematički izraz:
  - accuracy\_score: ~54.0
  - prosečan FuzzyWuzzy ratio: ~86
4. Za slike koje sadrže više zarotiranih matematičkih izraza:
  - accuracy\_score: ~44
  - prosečan FuzzyWuzzy ratio: ~82

## Problematične situacije

1. Često se desi da mreža pomeša simbol za vektorski proizvod i simbol 'x' što je i razumljivo jer su u pitanju veoma slični simboli

$$\begin{aligned} \text{Thuth expression} \\ 78pm5times47 \\ \text{Predicted expression} \\ 78pm5times47 \end{aligned}$$

2. Ponekad se desi da dva simbola ostanu spojena pa budu prepoznati kao jedan simbol

$$\begin{aligned} \text{Thuth expression} \\ 1pm20-173 \\ \text{Predicted expression} \\ 1pmh-173 \end{aligned}$$

3. U slučaju simbola *cos*, *sin*, *lim* i *log* ako pojedinačni karakteri nisu spojeni, mreži će biti prosleđeni pojedinačni karakteri za prepoznavanje umesto celog simbola pa neće biti dobro prepoznat ceo izraz. Prediktovani izraz je "c02(pi)=-1"

$$\cos(\pi) = -1$$

4. Ponekad se desi da, kod rotiranih matematičkih izraza, ugao rotacije ne bude dobro određen. Samim tim slika ne bude dobro ispravljena pa se ni prepoznavanje simbola ne izvrši uspešno. Prediktovani uzraz je: "x-,"

$$e^{-1}$$

5. Takođe se i kod slika sa više matematičkih izraza ponekad desi da rotirani izraz ne bude pravilno ispravljen. Prediktovani izraz je: "x=1z"

$$\begin{aligned} 3n^2 + 2n \\ x = \frac{1}{y} \\ 2e^1 - 3e - 2 - 3i = 0 \end{aligned}$$