**Universiteti i Prishtinës**

**“Hasan Prishtina”**



**Projekti 3**

**Lënda:** “Siguria në Internet”

**Tema:** Përdorimi i “webcrawling” për nxjerrjen e të dhënave

nga 1 webfaqe (Shembull: “Subreddits”)

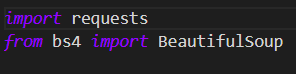
*Prof. Asoc. Blerim Rexha Punoi:*

*Msc. Arbnor Halili Ndriçim Halili*

*Prishtinë, 2021*

# Modulet e përdorura

Modulet e përdorura për këtë detyrë kanë qenë:



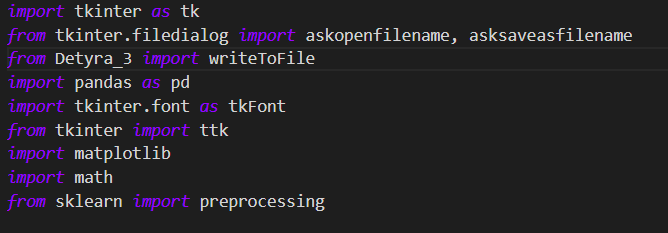


Figure Modulet

Shumica janë module ekzistuese, por disa, që kanë qenë të domosdoshme, sikur “requests”, nuk janë të gatshme.

# Aplikacioni



Figure 2 Aplikacioni

Për këtë detyrë kam vendosur që të marr të dhënat nga uebfaqja e njohur, “Reddit.com”, më saktë, nga (old.reddit.com).

Aplikacioni është i thjeshtë.

Në fillim ne mund të caktojmë një “subreddit” për të marrë vlerat, këtë e shkruajmë te fusha me “R/” dhe zgjedhim një “Subreddit” si dhe llojin e rendtitjes (“Hot”, “New”,”Rising”,”Top”,”Gilded”), sikurse në figurën më poshtë:



Figure 3 Aplikacioni

Më pas shtypim butonin “GO!”, i cili do të na nxjerri 11 rezultate si më poshtë:

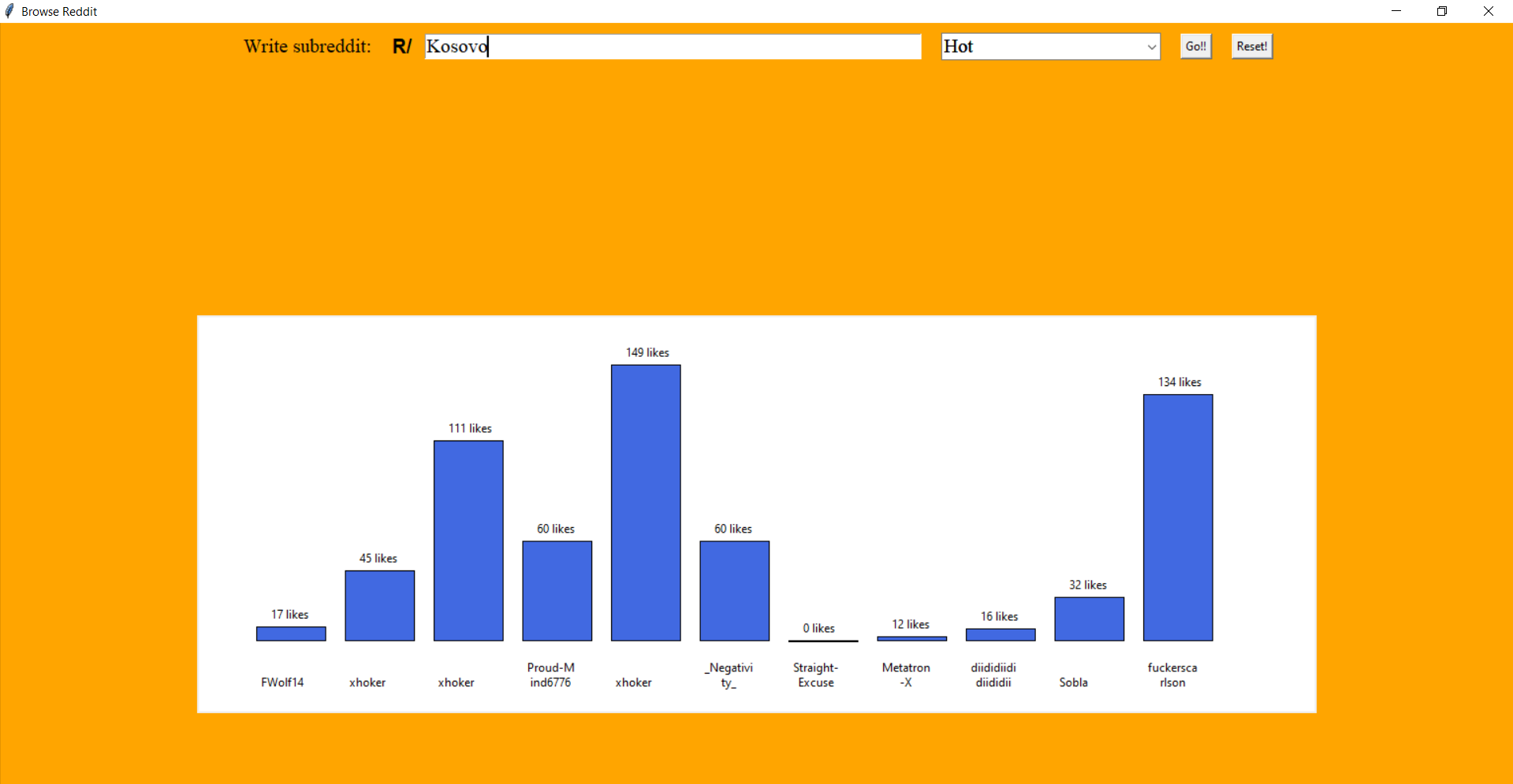


Figure 4 Rezultatet nga WebCrawling

Grafet përmbajnë të dhëna rreth numrit të pëlqimeve dhe poshtë tyre emri (“username”) i autorit të postimit. Grafet janë të renditura sipas çfarëdo që ka në “Combobox”, në këtë rast “Hot”.

Ne mund të shtypin njërin nga grafet për të parë titullin që e ka:

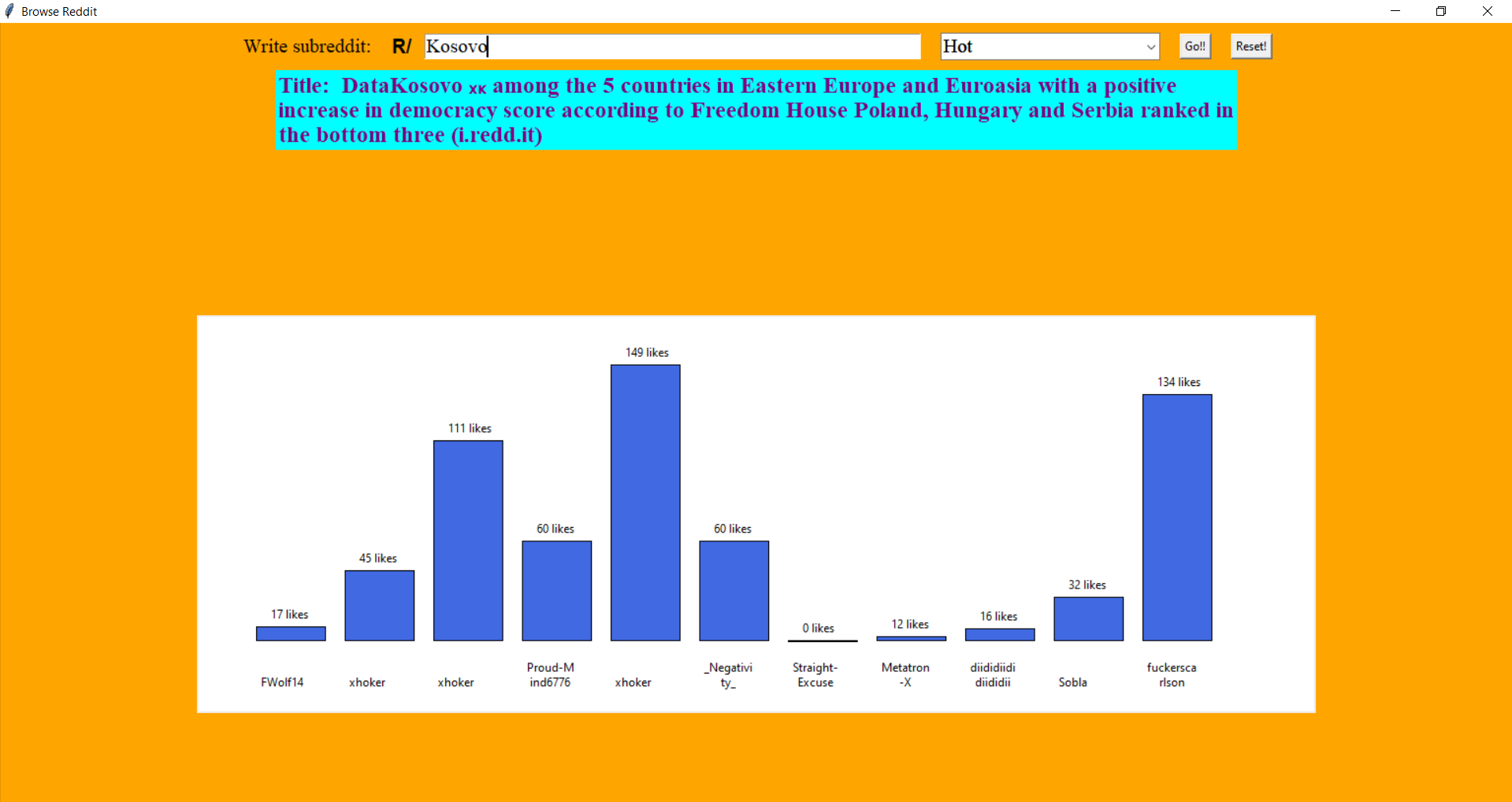


Figure 5 Grafet me tituj

Së fundmi ne mund të shtypim butonin “Reset” për të kthyer gjithçka sikurse në fillim.

Shënim: Nëse nuk jepet ndonjë “Subreddit” ose një lloj të renditjes, atëherë si vlera “default” do të merren “R/Kosovo” me renditje “Hot”

Shtesë: Nëse jepet ndonjë “Subreddit” që nuk ekziston ose nuk përdoruesi nuk ka qasje në “Reddit” ose në internet, atëherë do të shfaqet:

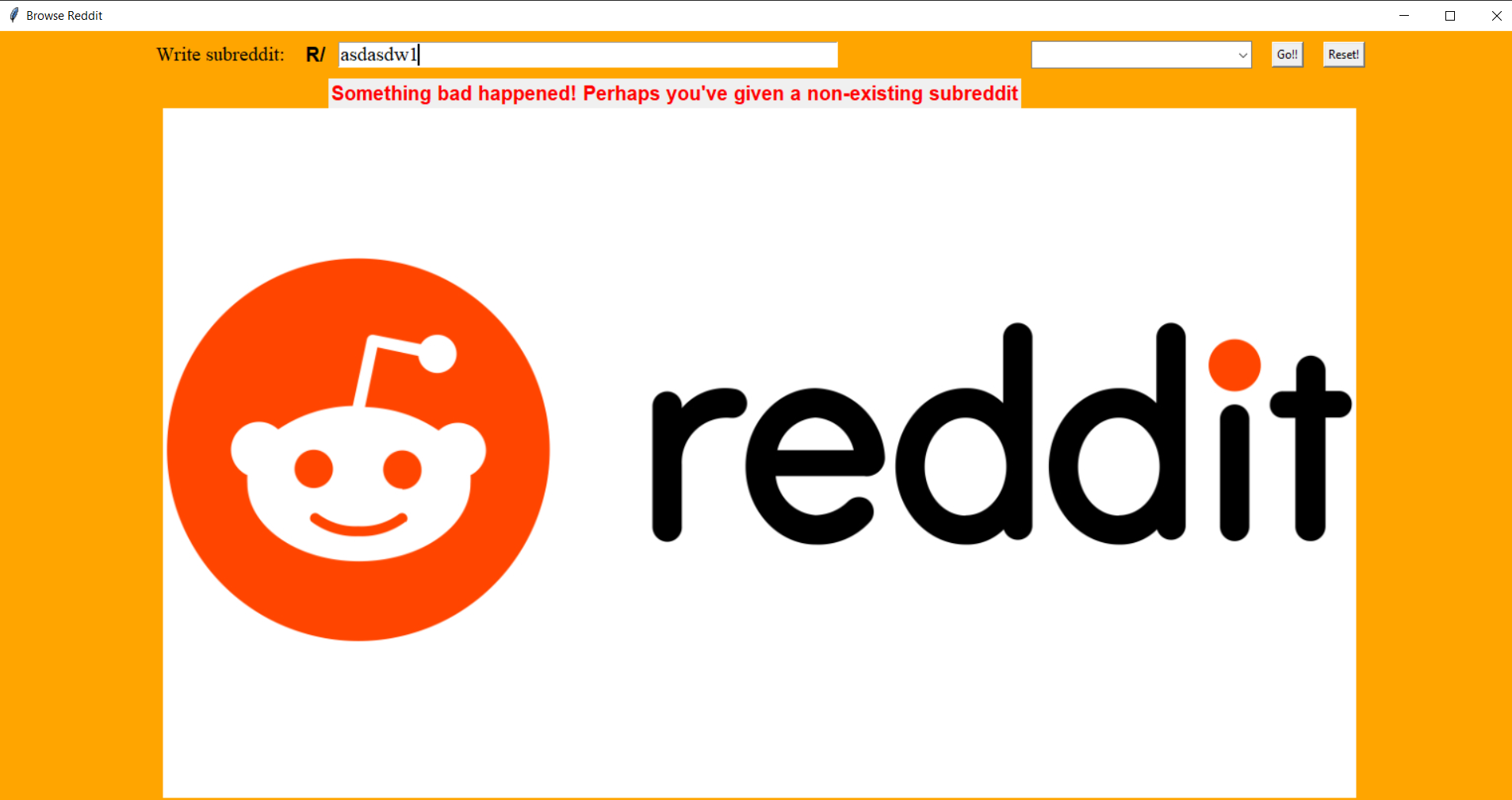


Figure 6 Gabime