

Predlog projekta iz SIAP-a

Ovaj dokument sadrži kratak opis onoga što su tema i definicija projekta, kao i motivacija za odabranu temu. Takođe. Naveden je i softver koji će biti korišćen za izradu projekta.

Tema projekta je sistem za analizu metoda klasterovanja (grupisanje po određenim osobinama) heroja u igrici League of Legends. League of Legends je video igra slična igri Dota. Meč igraju 10 igrača, koji su podeljeni u 2 tima. Svaki tim ima svoju bazu, i cilj je uništiti protivničku bazu. Timovi se međusobno bore tokom partije, i jedan od ciljeva je “ubiti” protivnika kako bi se dobili određeni resursi (“gold”, “experience”). Igrači tokom partije kupuju “iteme”, i tako biraju na koji način žele da razvijaju svog heroja. Svaki “item” daje neke dodatne osobine heroju i na osnovu tih osobina se može videti kako igrači generalno razvijaju heroje, što predstavlja dobre osobine za klasterovanje. Osobine će se posmatrati samo za mapu 5v5 (5 na 5).

Definicija projekta

Grupisanje heroja u igri na osnovu sličnih osobina, načina na koji se razvijaju, resursa koji se ulažu i njih i njihovih mogućnosti na osnovu istorije igranja tih heroja, koju omogućava API same igre. Ideja ovog projekta je da se na osnovu podataka klasteri jasno prepoznaju, i da predstavljaju kategorije heroja, npr “assasin”, “tank”, “support”, “marksman”, itd, kao i da se vidi koje kategorije su međusobno slične i koje su najpopularnije.

Pregled radova na sličnu temu

[1] Kevin Conley, Daniel Perry (2013) *How Does He Saw Me? A Recommendation Engine for Picking Heroes in Dota 2*

<http://cs229.stanford.edu/proj2013/PerryConley-HowDoesHeSawMeARecommendationEngineForPickingHeroesInDota2.pdf>

Tema rada: Sistem za preporuku heroja u igri Dota 2. Sistem uzima u obzir heroje koji su već izabrani u igri od strane saigrača i protivnika. Na osnovu toga traži se heroj koji će najbolje iskoristiti prednosti saigrača i mane protivnika i neutralisati prednosti protivnika i mane saigrača.

Skup podataka: Podaci su uzeti sa Valve's Steam web API-ja za 56691 mečeva između 5.10.2013. i 7.12.2013 i predstavljeni su u JSON formatu. Data set zadovoljava sledeće kriterijume:

- 1) Vrsta meča je *all pick*, *single draft*, *all random*, *random draft*, *captain's draft* ili *captain's mode*.
- 2) Uzeti su mečevi najboljih igrača (najboljih 8%)
- 3) Izbačeni su mečevi u kojoj je neki igrač napustio meč pre završetka.

Skup podataka sadrži informacije o herojima koji su izabrani u meču, njihovim rezultatima i o timu koji je pobedio. 90% skupa podataka je korišćen za trening set, a 10% za testiranje.

Korišćeni algoritmi: *Logistic regression* I *K-nearest-neighbour*.

[2] Nicholas Kinkade, Kyung yul Kevin Lim (2015) *DOTA 2 Win Prediction*

<http://cseweb.ucsd.edu/~jmcauley/cse255/reports/fa15/018.pdf>

Tema rada: Model za predikciju ishoda meča u Dota 2 igri na osnovu dva prediktora. Prvi prediktor uzima u obzir sve podatke vezane za meč, dok drugi prediktor uzima u obzir samo kompozicije timova. Rađena je eksplorativna analiza I računata je korelacija heroja u okviru tima, kao i međusobni odnos heroja suprotstavljenih timova. Prvi prediktor je služio za validaciju njihovog rada.

Skup podataka: Podaci su uzeti sa *Valve's Steam web API*-ja za 62000 mečeva između 20.11.2015. I 22.11.2015 I predstavljeni su u *JSON* formatu. Data set zadovoljava sledeće kriterijume:

- 1) Vrsta igre je 5v5 (5 na 5)
- 2) Uzeti su mečevi najboljih igrača (najboljih 8%)
- 3) Izbačeni su mečevi u kojoj je neki igrač napustio meč pre završetka
- 4) Izbačeni su mečevi koji su trajali manje od 10 minuta.

Skup podataka sadrži informacije o pobjedniku, trajanju meča, I za svakog igrača *Gold Per Minute*, *Kills*, *Assists*, *Deaths*.

Korišćeni podaci: Prvi prediktor je koristio *random forests*, dok je drugi koristio *logistic regression*

Rezultati: Prvi prediktor je imao gotovo savršene rezultata, dok je drugi imao oko 73% tačnosti.

Skup podataka

Za ovu temu će se koristiti skup podataka sa API-ja same igre - <https://developer.riotgames.com/>. Skup podataka je predstavljen u "json" formatu. Ti podaci predstavljaju istoriju mečeva I ne sadrže sve potrebne informacije. Podaci će zatim biti obrađeni na više načina, kako bi se mogli uporediti rezultati. Na primer, iz podataka se može dobiti informacija o tome koje "iteme" je igrač kupio tokom partije, ali ne I kakve koristi igrač ima od takvih itema, tj kakve osobine mu daje. Te informacije se mogu dobiti posebno za svaki "item", tako da će biti potrebno posebno izvući te podatke. Na slikama u prilogu se nalazi prikaz dela skupa podataka koji će se koristiti za izradu projekta.

```

object {1}
  matches [100]
    0 {15}
      matchId : 2054994244
      region : NA
      platformId : NA1
      matchMode : CLASSIC
      matchType : MATCHED_GAME
      matchCreation : 1451603074200
      matchDuration : 3067
      queueType : RANKED_SOLO_5x5
      mapId : 11
      season : PRESEASON2016
      matchVersion : 5.24.0.256
      participants [10]
      participantIdentities [10]
      teams [2]
      timeline {2}
    1 {15}

```

```

participants [10]
  0 {10}
    teamId : 100
    spell1Id : 4
    spell2Id : 7
    championId : 236
    highestAchievedSeasonTier : UNRANKED
    timeline {6}
    masteries [10]
    stats {58}
    participantId : 1
    runes [4]
  1 {10}

```

```

stats {58}
  winner : false
  champLevel : 18
  item0 : 3034
  item1 : 3031
  item2 : 3139
  item3 : 1301
  item4 : 3508
  item5 : 3026
  item6 : 3363
  kills : 6
  doubleKills : 1
  tripleKills : 1
  quadraKills : 0
  pentaKills : 0
  unrealKills : 0
  largestKillingSpree : 4
  deaths : 13
  assists : 16
  totalDamageDealt : 278697
  totalDamageDealtToChampions : 41886
  totalDamageTaken : 47802
  largestCriticalStrike : 1138
  totalHeal : 7118
  minionsKilled : 217
  neutralMinionsKilled : 46
  neutralMinionsKilledTeamJungle : 32
  neutralMinionsKilledEnemyJungle : 14
  goldEarned : 17977
  goldSpent : 17550
  combatPlayerScore : 0
  objectivePlayerScore : 0

```

```

totalPlayerScore : 0
totalScoreRank : 0
magicDamageDealtToChampions : 3445
physicalDamageDealtToChampions : 37598
trueDamageDealtToChampions : 841
visionWardsBoughtInGame : 0
sightWardsBoughtInGame : 0
magicDamageDealt : 12664
physicalDamageDealt : 263863
trueDamageDealt : 2169
magicDamageTaken : 16501
physicalDamageTaken : 29673
trueDamageTaken : 1628
firstBloodKill : false
firstBloodAssist : false
firstTowerKill : false
firstTowerAssist : false
firstInhibitorKill : false
firstInhibitorAssist : false
inhibitorKills : 0
towerKills : 1
wardsPlaced : 20
wardsKilled : 1
largestMultiKill : 3
killingSprees : 1
totalUnitsHealed : 3
totalTimeCrowdControlDealt : 184
participantId : 1

```

Odabir odgovarajućih metoda I algoritama

Koristiće se razne metode klasterovanja I poređiće se rezultati. Neke od metoda koje će biti korišćene su *K-Means*, *K-Modes*, *Mean-shift*, *Spectral clustering*, *DBSCAN*. Izbor najprikladnijih algoritama zavisi od načina odabira I predstave podataka, što će se menjati tokom izrade projekta.

Pored toga, radiće se redukcija dimenzionalnosti I biće poređeni rezultati. Redukcija dimenzionalnosti će se pokušati ostvariti PCA algoritmom, ukoliko rezultati ne budu dobri, biće isprobani drugi algoritmi, greedy pristupi ili nelinearno mapiranje.

Softver

Za izradu projekta će se koristiti *Python* programski jezik I *SciKit Learn*.

Plan rada

- ⑩ Prikupljanje podataka sa <https://developer.riotgames.com/>
- ⑩ Transformacija podataka, dodavanje I izbacivanje atributa ručno
- ⑩ Eventualno automatski transformisanje kategoričkih atributa
- ⑩ Primena metoda klasterovanja
- ⑩ Vizuelizacija I upoređivanje rezultata, PCA
- ⑩ Analiza rezultata će biti rađena eksplorativnom analizom, I nekom od metoda interni indeks, pri čemu će se uporediti rezultati različitih algoritama. Za pocetak će biti korišćen *Silhouette Value* metod.

Tim

Milan Keča, E2 67/2016