**Veri Analizi**

Veri seti, Dota2 ismindeki internet üzerinden oynana bir oyunun maçlarını içermektedir. Veri setinde toplam 102.944 örnek bulunmaktadır. Veri setinde iki sınıftan oluşmaktadır. Bu sınıflar maçın hangi takım tarafından kazanıldığını belirtmektedir. Bu kategorik değerler, 1 ve -1 olarak gösterilmektedir. Verilerin 48.660 tanesi -1 ile etiketlenmişken 54.284 tanesi ise 1 ile etiketlenmiştir.

Verisetinde 116 tane özellik, 117 sütun bulunmaktadır. Baştan sonra doğru sütunlar: maçı hangi takımın kazandığını (sınıf bilgisi), cluster id, oyun modu, oyun tipi ve geri kalan 113 sütun oyun içerisindeki karakterlerin maç esnasında hangi takımda olduklarını göstermektedir. 113 tane karakter sütunlarının bir maçta (örnekte) 10 tanesi dolu olabilmektedir. Ait oldukları takımlara göre -1 ve 1 değerlerini almaktadırlar. Eğer herhangi bir takımda değilse o karakter değeri 0 olmaktadır.

Csv formatındaki dosyayı arff formatına çevirirken sütunların isimlerine ihtiyaç duyuldu. Sitede belirtilen (<https://archive.ics.uci.edu/ml/datasets/Dota2+Games+Results>) mapping kullanılarak sütunlardaki herolar ilgili şekilde isimlendirildi. Bir tane hero için mapping eksikti. Bunun için veri setinin oluşturulması için hazırlanan kod incelenerek (<https://gist.github.com/da-steve101/1a7ae319448db431715bd75391a66e1b>) veri çekmek için OpenDota API’nın kullanıldığı tespit edildi. OpenDota API’ya bakılarak eksik sütun bilgisi tamamlandı. Sütunların değerlerinin ne anlama geldiği ise veri seti dökümanında bulunmadığından ötürü gene veri seti için hazırlanan kod incelenerek kullanılan wrapper’ın dokümantasyonundan (<https://github.com/joshuaduffy/dota2api/blob/master/docs/source/responses.rst>) alındı.

“game\_mode” sütunun alabileceği değerler 0’dan 22’ye kadar sırasıyla Unknown, All Pick, Captain's Mode, Random Draft, Single Draft, All Random, Intro, Diretide, Reverse Captain's Mode, The Greeviling, Tutorial, Mid Only, Least Played, New Player Pool, Compendium Matchmaking, Custom, Captains Draft, Balanced Draft, Ability Draft ,Event, All Random Death Match, Solo Mid 1 vs 1, Ranked All Pick’e denk gelmektedir.

“game\_type” sütunun alabileceği değerler -1’den 8’e kadar sırasıyla Invalid, Public matchmaking, Practice, Tournament, Tutorial, Co-op with AI, Team match, Solo queue, Ranked matchmaking, Solo Mid 1 vs 1’e denk gelmektedir.

Weka içerisinde bulunan InfoGainAttributeEval algoritması kullanılarak veri seti içerisindeki ayırt edici özellikler belirlendi. İlk beş belirleyici özellik: enchantress, mirana, life\_stealer, windrunner, abyssal\_underlord.