

Karar Ağacı İle Çalışan Üretkenliği Tahminleme

...

Kardelen E.

İçerik

1. Giriş
2. Ağaç Modeli
3. Model Açıklaması
4. Örnek Senaryolar
5. Model Kalitesi
6. Sonuç

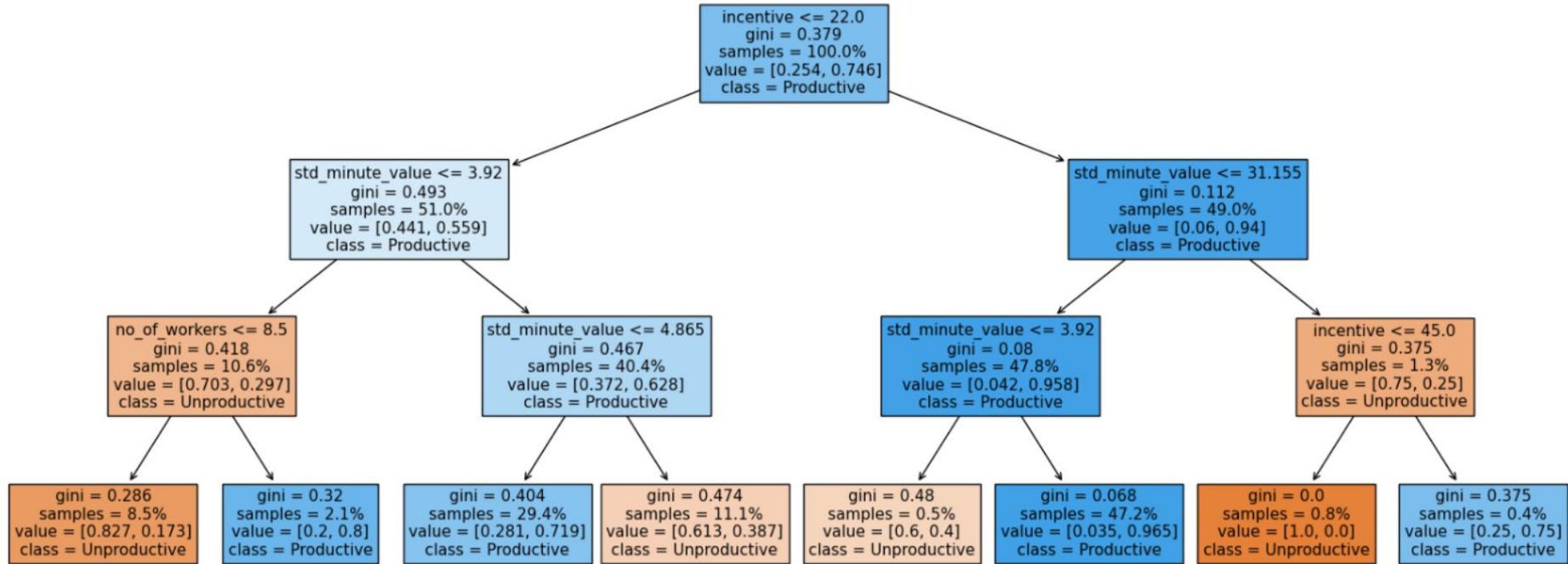
1. Giriş

Karar ağacı, teknoloji bilgisi yüksek olmayan insanlar için anlaması kolay bir makine öğrenme tipidir. Bu projede tekstil endüstrisinden bir veriyi inceleyeceğiz ve karar ağacı kullanarak çalışan üretkenliğini tahminleyeceğiz.

Tekstil endüstrisi yüksek el emeği isteyen bir endüstri olduğu için çalışan üretkenliğini tahminlemek şirketteki karar alıcılar için büyük öneme sahiptir.

	date	quarter	department	day	team	targeted_productivity	smv	wip	over_time	incentive	idle_time	idle_men	no_of_style_change	no_of_workers	actual_productivity
0	1/1/2015	Quarter1	sweing	Thursday	8	0.80	26.16	1108.0	7080	98	0.0	0	0	59.0	0.940725
1	1/1/2015	Quarter1	finishing	Thursday	1	0.75	3.94	NaN	960	0	0.0	0	0	8.0	0.886500
2	1/1/2015	Quarter1	sweing	Thursday	11	0.80	11.41	968.0	3660	50	0.0	0	0	30.5	0.800570
3	1/1/2015	Quarter1	sweing	Thursday	12	0.80	11.41	968.0	3660	50	0.0	0	0	30.5	0.800570
4	1/1/2015	Quarter1	sweing	Thursday	6	0.80	25.90	1170.0	1920	50	0.0	0	0	56.0	0.800382

2. Ağaç Modeli



3. Model Açıklaması

(Modelin teknik detayları ipynb dosyasında mevcut)

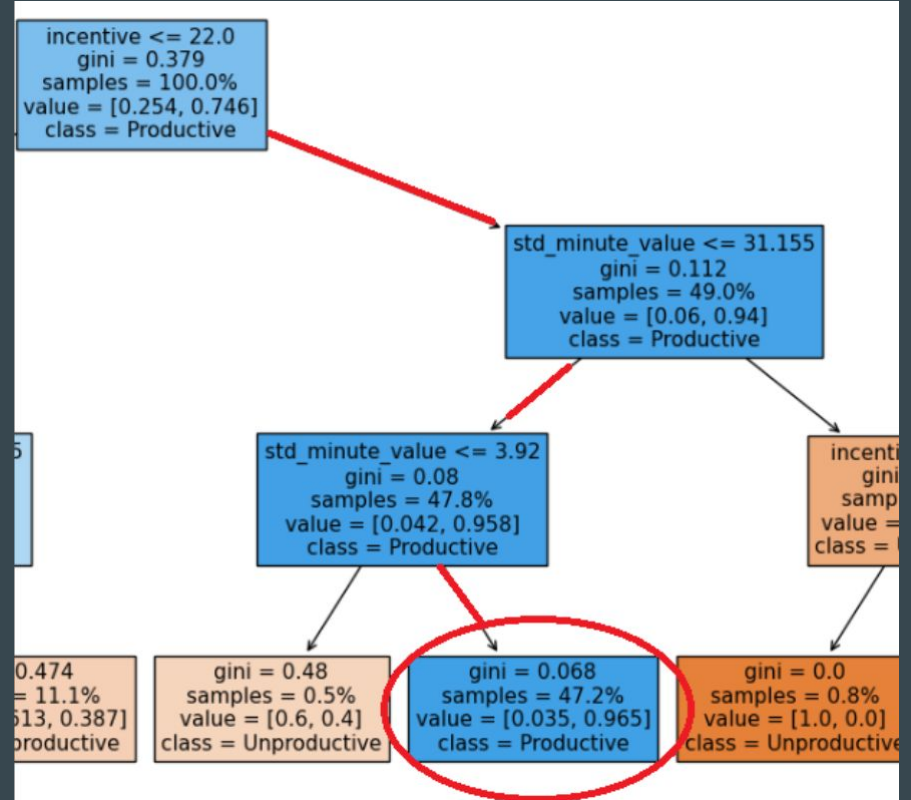
Model, en önemli özelliklerin incentive (teşvik) ve standard minute value (standart dakika değeri) olduğunda karar kıldı. Karar ağacını sadece 3 dala sahip olmakla sınırlamasaydık diğer özelliklerin de kullanılacağını gözden kaçırmamak gerekir.

Ağacı açıklamanın en iyi yolu örnek senaryolar üretmektir.

4.1 Örnek Senaryolar

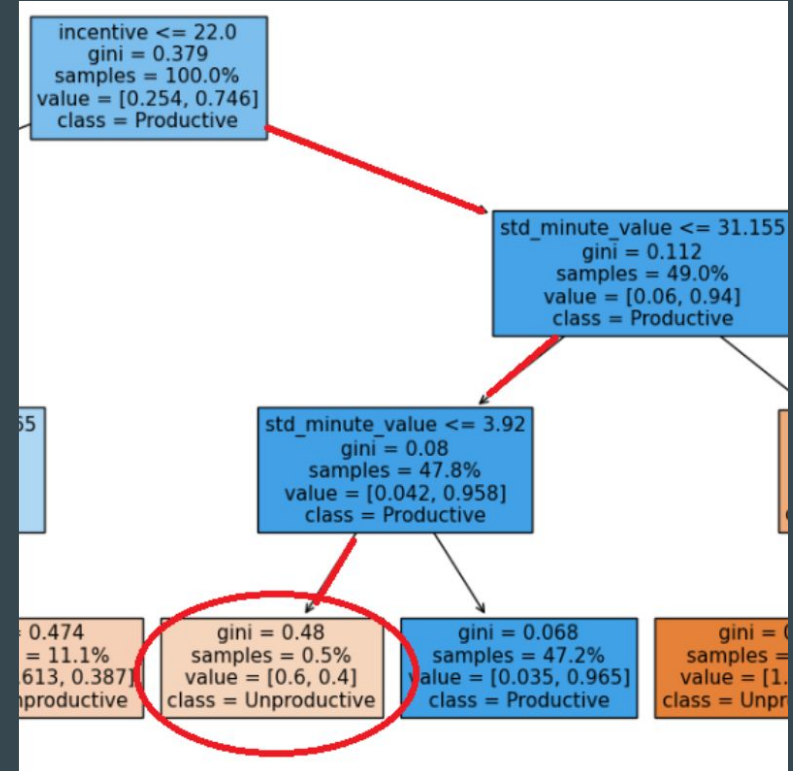
“Incentive”in 24 ve “standard minute value”nun 5 olduğu bir senaryo düşünelim. Modelimiz o gün için üretkenlik hedefine ulaşılacağını tahminler miydi?

“Incentive” 22’den büyük olduğu için “24 <= 22” YANLIŞ olarak döner ve bu yüzden sağa gideriz (**YANLIŞ ayrımlar her zaman sağa gider**). “Standard minute value” 31,155’ten az olduğu için “5 <= 31,155” DOĞRU olarak döner ve rotamıza sol yönde devam ederiz (**DOĞRU ayrımlar her zaman sola gider**). Son ayrımda “standard minute value” 3,92’den büyük olduğu için YANLIŞ rotasından devam ederiz ve “Productive” (Üretken) sınıfına erişiriz. Bu gün üretken bir gün olacaktır.



4.2 Örnek Senaryolar

Farklı bir senaryo deneyelim. Bu sefer “incentive” 24 olsun ve “standard minute value” 3 olsun. “Incentive” yine 22’den büyük olacağı için YANLIŞ rotasından ilerleriz. “Standard minute value” 3 olduğu için ve 31,155’ten küçük olduğu için DOĞRU rotasından devam ederiz. Son ayırmada “Standard minute value” bu senaryoda 3.92’den küçük olduğu için tekrar DOĞRU rotasından ilerleriz ve rotamız bu sefer “Unproductive” (Üretken olmayan) sınıfında tamamlanır. Bu gün üretken olmayacaktır. Ağaç modeli veri setindeki örneklerin %0,5’inin burada toplandığı bilgisini de bize veriyor.



5. Model Kalitesi

Modelimizin kalitesini görmek için bazı istatistiksel sayılara bakalım.

- Precision (PPV, Kesinlik) şöyle hesaplanır:

$$PPV = \frac{\text{Gerçek pozitif sayısı}}{\text{Gerçek pozitif sayısı} + \text{Yanlış pozitif sayısı}}$$

Modelimizin kesinlik değeri %88.

- Recall (NPV, Duyarlılık) kesinlik ile aynı şekilde hesaplanır, ancak negatif sayıları kullanılır. Modelimizde bu sayı %93.
- F-skor kesinlik ve duyarlılık kullanarak hesaplanır. Modelimizde bu sayı %90.

Tüm bu bulgulardan modelimizin kalitesinin yüksek olduğu söylenebilir.

6. Sonuç

Tekstil endüstrisinden aldığımız veriyi temizledik ve üretkenlik hedefine ulaşıp ulaşılmayacağını gösteren bir karar ağacı modeli ürettik.

Modelimizin birtakım istatistiksel özelliklerine baktık ve kalitesinin yüksek olduğunu gördük. Hayali veriler kullanarak ve model üzerinde ilerleyerek bu potansiyel günlerde üretkenlik hedefine ulaşıp ulaşılmayacağını hesapladık.