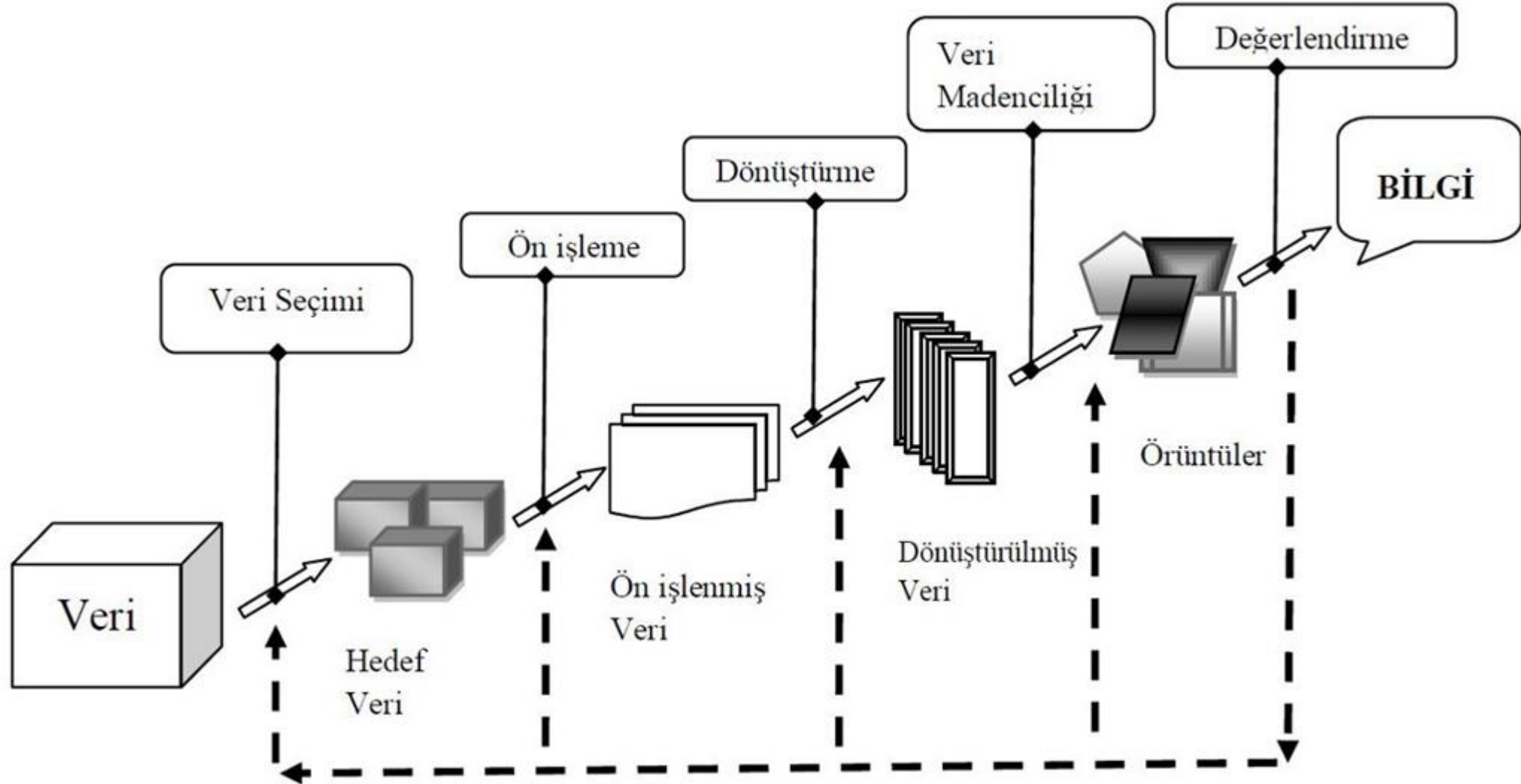
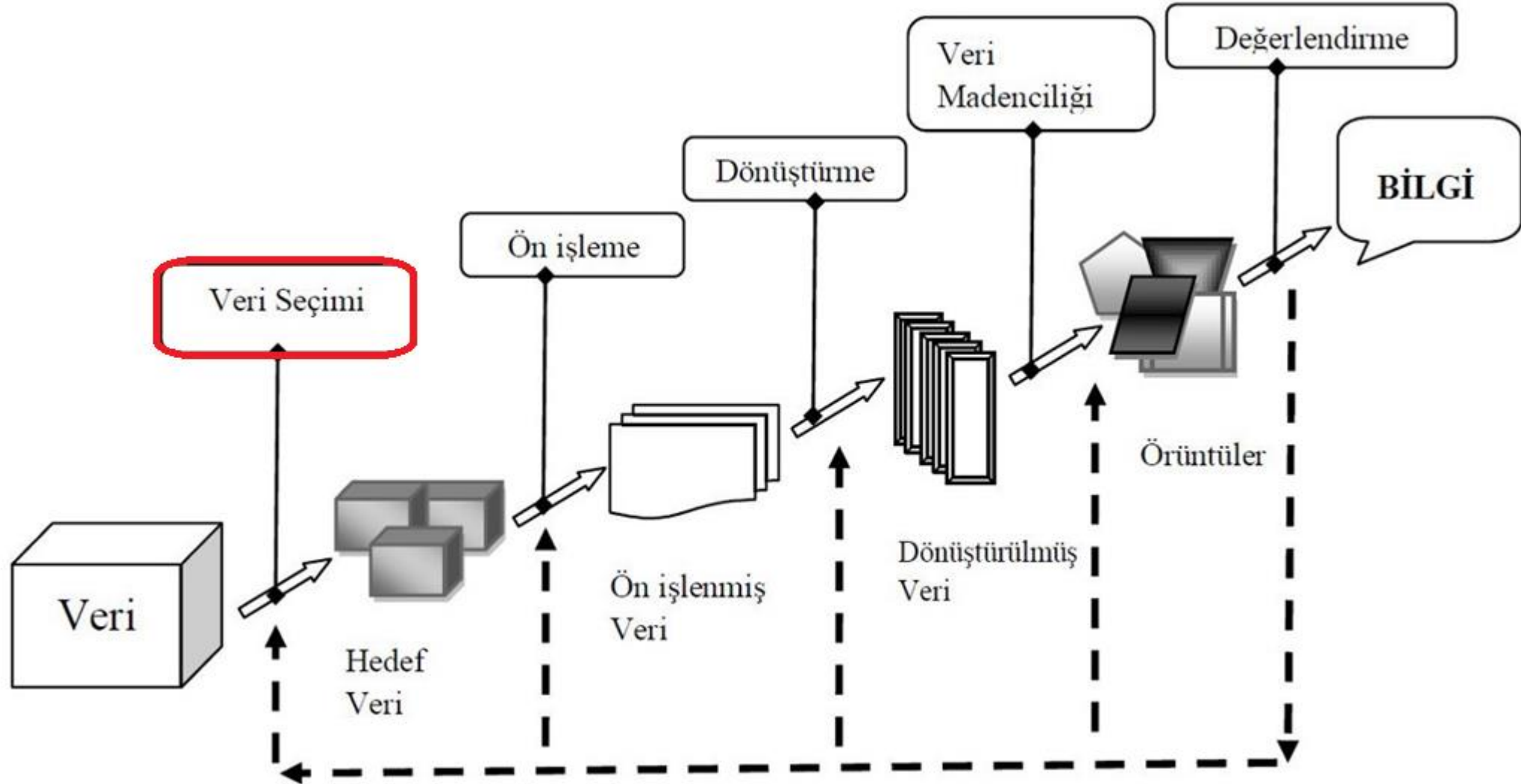


# Veri Madenciliği Süreci

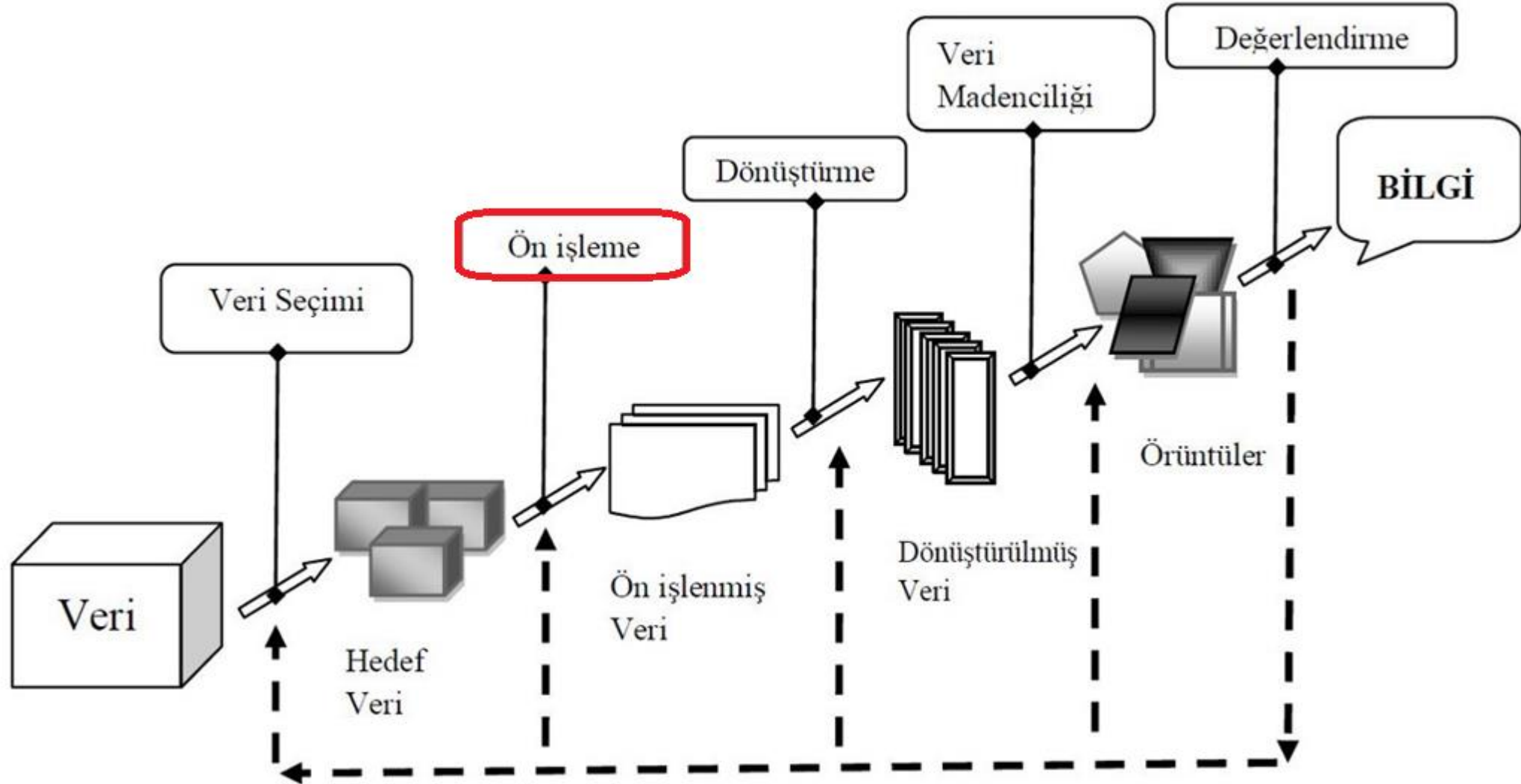
# Veri Madenciliği Süreci



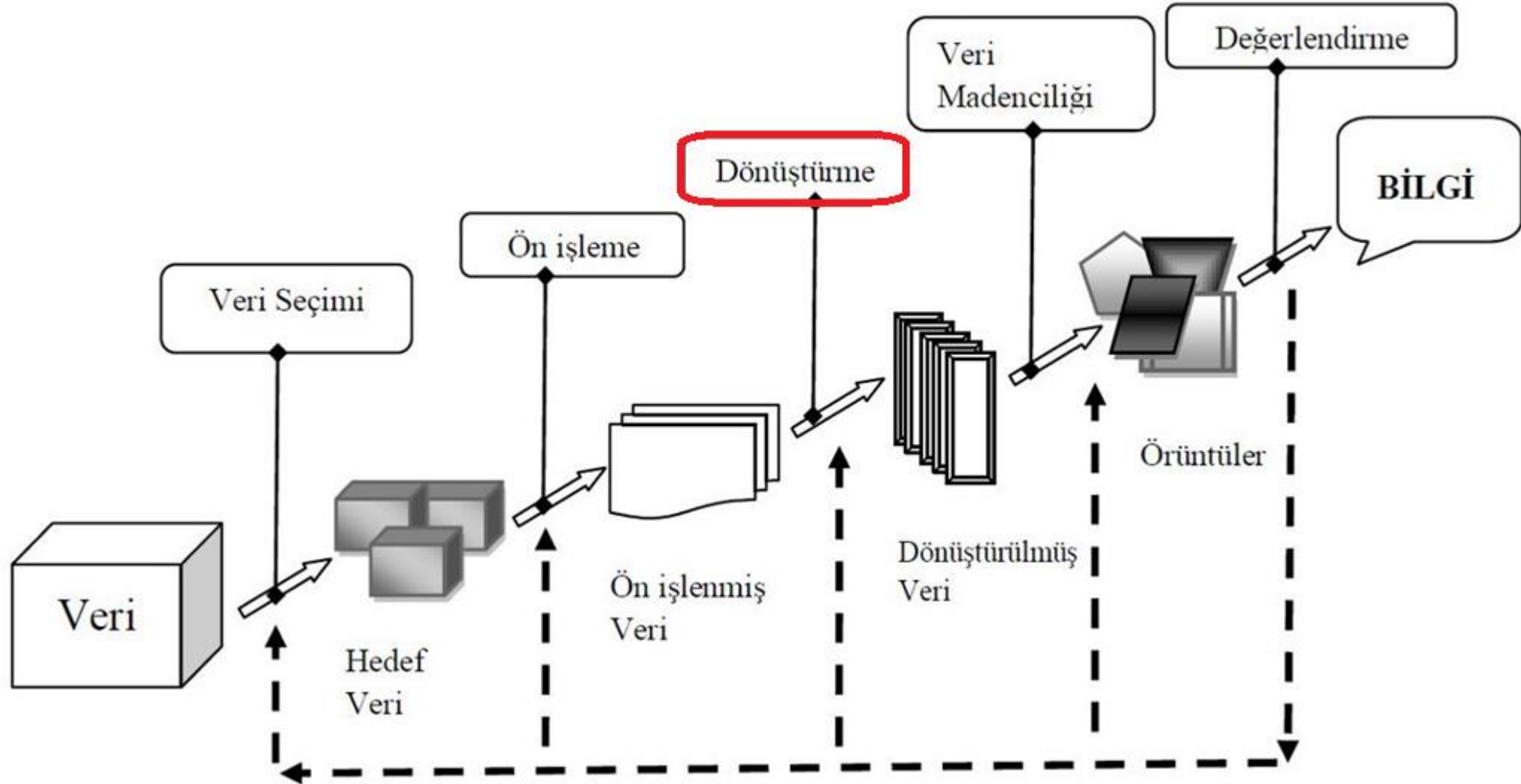
# Veri Madenciliği Süreci



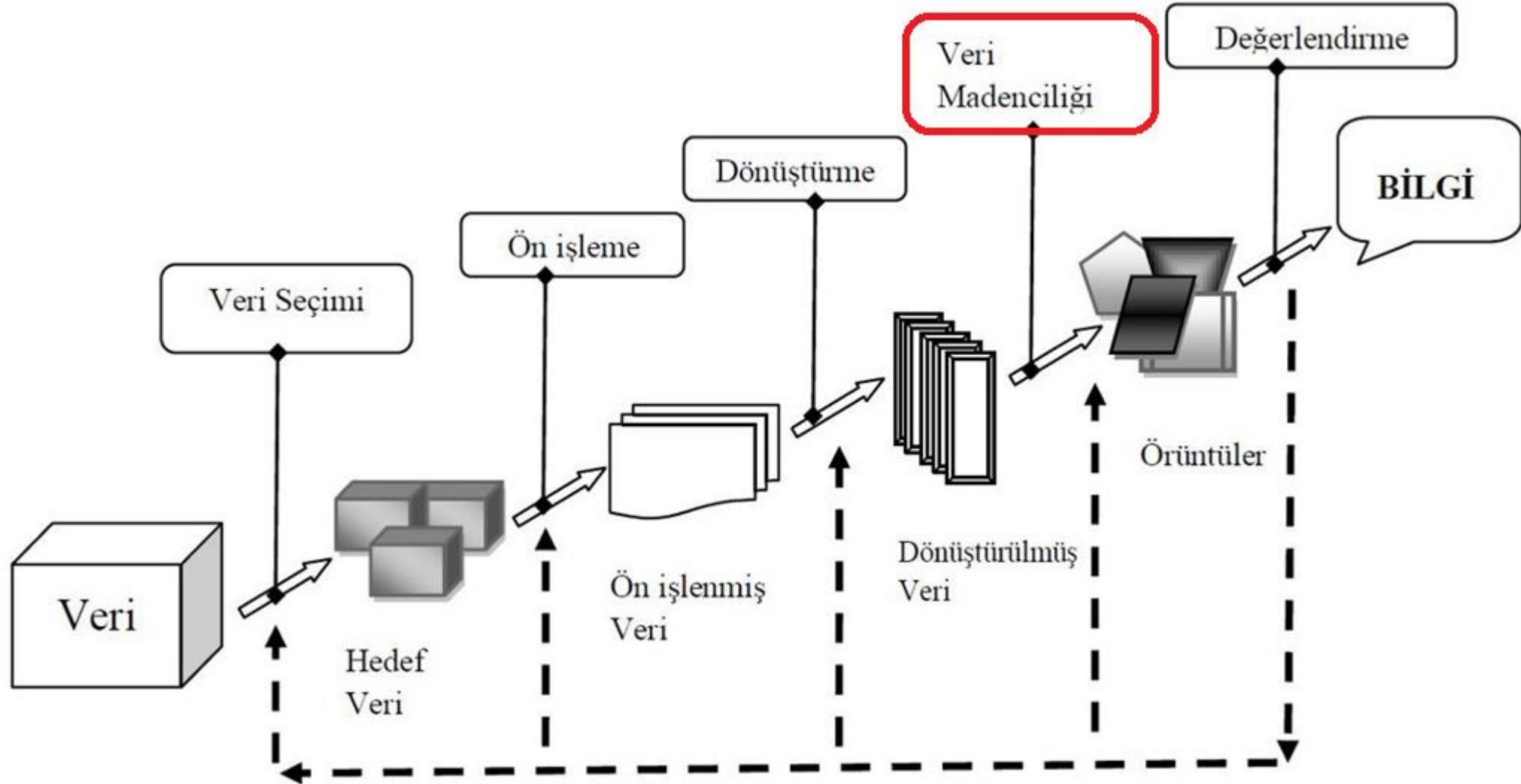
# Veri Madenciliği Süreci



# Veri Madenciliği Süreci

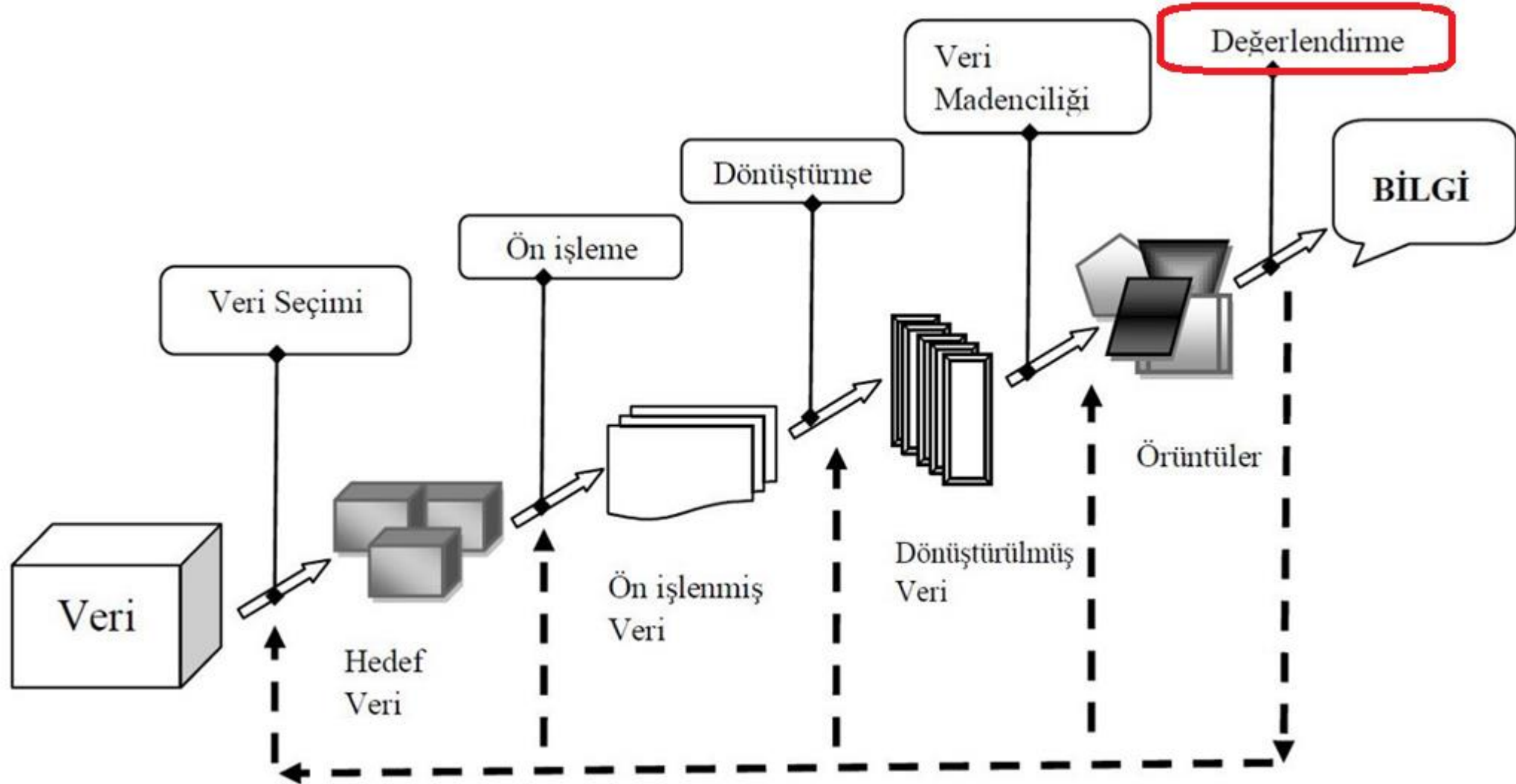


# Veri Madenciliği Süreci





# Veri Madenciliği Süreci



# Veri Seçimi

- Veri Madenciliği uygulanacak konu ile ilgili veriler belirlenmelidir.
  - Son 5 Yılın verileri
  - Öğrencilerin not ortalamaları
  - Sadece Meslek Dersleri
  - ...



# Ön İşleme

- **Veri Temizleme:** Eksik ve Uygun olmayan verilerin temizlenmesi, düzenlenmesi gerekir.

**Gürültü**



# Ön İşleme (Veri Temizleme)

- Eksik değerler atılabilir,
- Genel bir sabit kullanılabilir,
- Ortalama değer alınabilir,
- Her sınıfa ait ortalama alınabilir,
- Tahmin modeli oluşturulabilir.

Ay	Sıcaklık	Nem	Basınç
Ocak	2,7	45	917
Şubat	3,9	NULL	910
Mart	10,6	51	905
Nisan	17,5	49	927
Mayıs	NULL	46	952
Haziran	28,7	52	NULL
Temmuz	30,2	51	923
Ağustos	33,5	NULL	890
Eylül	24,6	43	895
...	...	...	...

# Ön İşleme

- **Veri Bütünleştirme:** Farklı veri tabanlarından ya da veri kaynaklarından elde edilen verilerin birlikte değerlendirmeye alınabilmesi için bütünleştirme gerekmektedir.

Ay	Sıcaklık (°C)	Nem	Basınç
Ocak	2,7	45	917
Şubat	3,9	48	910
Mart	10,6	51	905
Nisan	17,5	49	927
Mayıs	23,3	46	952

Ay	Sıcaklık (F)	Nem	Basınç
Ocak	33,8	45	917
Şubat	42,9	44	910
Mart	50,7	51	905
Nisan	57,8	49	927
Mayıs	65,9	46	952

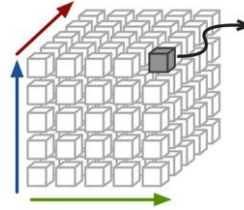
# Ön İşleme

- **Veri İndirgeme:** Eğer çözümlmeden elde edilecek sonucun değişmeyeceğine inanılıyorsa veri sayısı ya da değişkenlerin sayısı azaltılabilir.

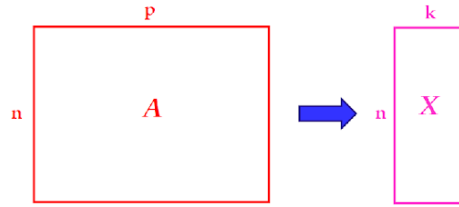


# Ön İşleme (Veri İndirgeme)

- Veri birleştirme veya veri küpü



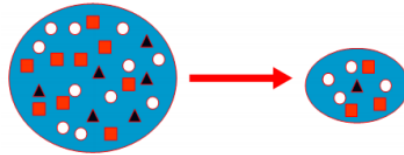
- Boyut indirgeme



- Veri sıkıştırma



- Örnekleme



- Genelleme



# Veri Dönüştürme

- Değişkenlerin arasındaki ortalama ve varyanslarının farklı olmaları sonucu büyük ölçüde olumsuz etkileyebilir.

Ay	Sıcaklık (°C)	Nem	Basınç
Ocak	2,7	45	917
Şubat	3,9	48	910
Mart	10,6	51	905
Nisan	17,5	49	927
Mayıs	23,3	46	952



Ay	Sıcaklık (°C)	Nem	Basınç
Ocak	0,0	0,0	0,3
Şubat	0,1	0,5	0,1
Mart	0,4	1,0	0,0
Nisan	0,7	0,7	0,5
Mayıs	1,0	0,2	1,0

# Veri Madenciliđi Uygulama

- Veri hazır hale getirildikten sonra konuyla ilgili veri madenciliđi algoritmaları uygulanır.





# Sunum ve Değerlendirme

- Veri madenciliği algoritması uygulandıktan sonra, sonuçlar düzenlenerek ilgili yerlere sunulur. Sonuçlar çoğu kez grafiklerle desteklenir.

