Bu yazının amacı, bilgisayar bilimlerinde özellikle veri madenciliği (data mining) konularında eksik verilerle karşılaşılması halinde bir çözüm olarak bu eksik verilerin töhmet edilmesi (yerine uygun verilerin üretilmesi, imputation) yöntemini açıklamaktır.

Töhmet, sözlükte olmayan birşeyin yüklenmesi anlamındadır. Örneğin olmayan bir suçun birisine yüklenmesine töhmet altında bırakmak denilebilir. Bu anlamda vir veri kümesi (data set) üzerinde çalışılırken bazı sebeplerden dolayı verilerin eksik olması halinde bu verilerin uygun başka sayılarla tamamlanması sağlanabilir.

Genelde verinin hatalı okunması, veri kaynağında yaşanan bozulma gibi sorunlar veya bazı verilere erişim zorluğu eksik verilere sebep olabilmektedir.

Bu verilerin eksik olması durumu genelde sorunlara sebep olur. Örneğin çoğu hazır istatistik paketleri (SAS, SPSS, Weka veya r-project gibi) bu tip durumlarda sorunlar yaşamaktadır. Gerçi bu paketlerin ücretli ve gelişmiş olanlarının çoğunda (SAS veya SPSS gibi) töhmet modülleri (imputation) bulunmakta ve bu işi otomatik olarak yapabilmektedirler ancak bu yazı kapsamında gerek bu modüllerin nasıl çalıştığını anlamaya çalışacağız gerekse bu işlemi elle yapmak istediğimizde nasıl müdahale etmemiz gerektiğini açıklamaya çalışacağız.

Töhmet yöntemleri duruma ve beklentilerimize göre çeşitlilik arz eder ve halen üzerinde çalışılmaktadır. Bunlardan çok bilinen bazılarını saymamız gerekirse: