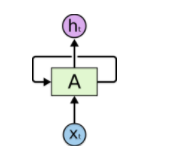
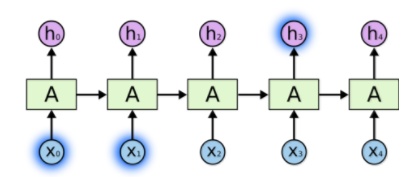
IV. Aşama

Bu hafta veri seti ile ilgili tüm düzenlemeler bitirildikten sonra veri madenciliği ve derin öğrenme algoritmaları üzerine bazı çalışmalar yapıldı. Yapılan çalışmalarda yinelenen sinir ağları(RNN), yapısal olarak bu çalışmaya daha uygun olacağı düşünülmüştür. Yinelenen sinir ağı yapısının seçimindeki en kritik yapı şu olmuştur:



Tekrarlayan sinir ağları da döngüler vardır. Yukarıdaki diyagramda, “A” girişi (xt) var ve çıkış değeri (ht) var. Bir döngü bir adımdan diğer sinir ağına bilgi gitmesine izin verir.

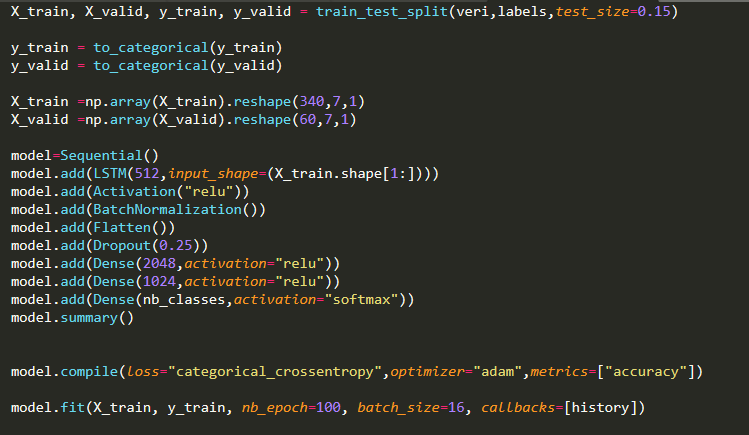
Bu durum kabaca bir örnek vermek gerekirse; Kurulan model “SO2” özelliğinin “Trafik\_Indeks” özelliğine etkisini öğrenirken, bu öğrenilmiş yapının “PM2.5” özelliğinin “Trafik\_Indeks” özelliğine etkisi öğrenilirken “SO2” özelliğinin öğrenildiği verileri ile kullanılması gibi tarif edilebilir.



Bu diyagramda döngü zincirinin açık olarak görülmesi özelliklerin birbiri arasındaki ilişkisini daha iyi açıklayacaktır.

Özetle açıklamak gerekirse bir özelliğin diğer özelliklerin eğitimi sırasında arasında bilgi akışını gerçekleştirmesi için bir RNN türü olan LSTM seçildi.

Tüm bu etkenler sonucunda veri seti üzerinde örnek bir LSTM modeli uygulandı. Model için kullanılan kod bloğu şu şekilde:



Bu model ile başarım ve kayıp oranları test edildi sonuçlarda ciddi bir kayıp olduğu için optimize çalışmaları yapılması gerekiyor. LSTM haricinde diğer algoritmalar ile başarım oranları test edilmeye devam ediliyor. En uygun ve en iyi öğrenme için optimize etme çalışmalarına başlandı.

**\*Diğer bir seçenek olarak RapidMiner ortamında çalışmalar yapılacak ve daha geçerli bilgiler elde edilmesi durumunda RapidMinerda oluşturulmuş olan model üzerinde çalışılacak.**