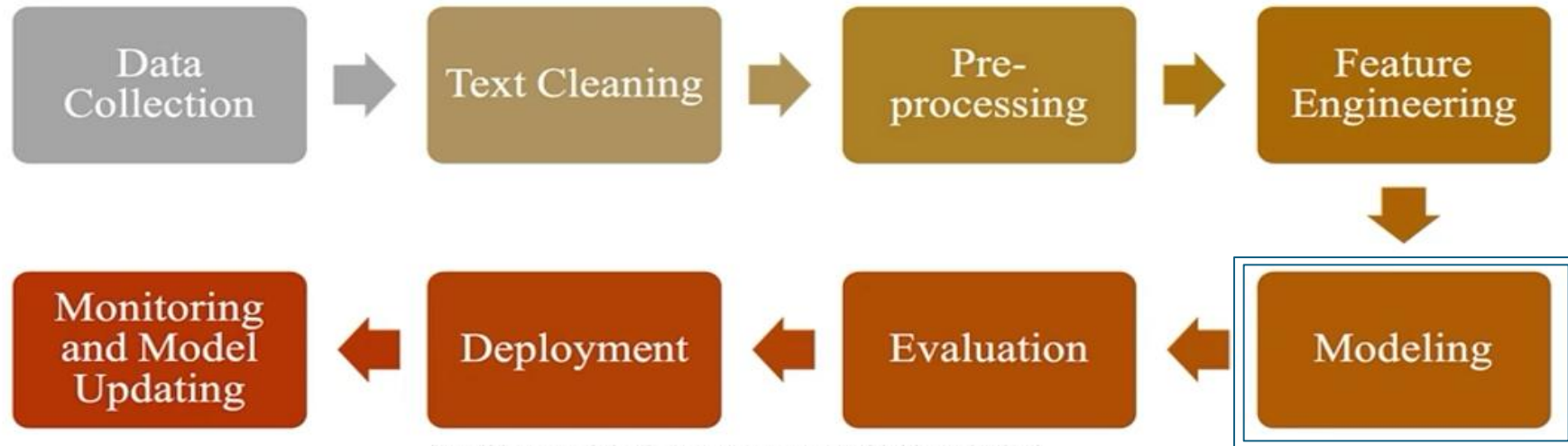


Olasılıksal Dil Modelleri

- Olasılıksal dil modelleri, dilin yapısını ve düzenini anlamak ve modellemek için kullanılan istatistiksel yöntemlerdir.

NLP Pipeline



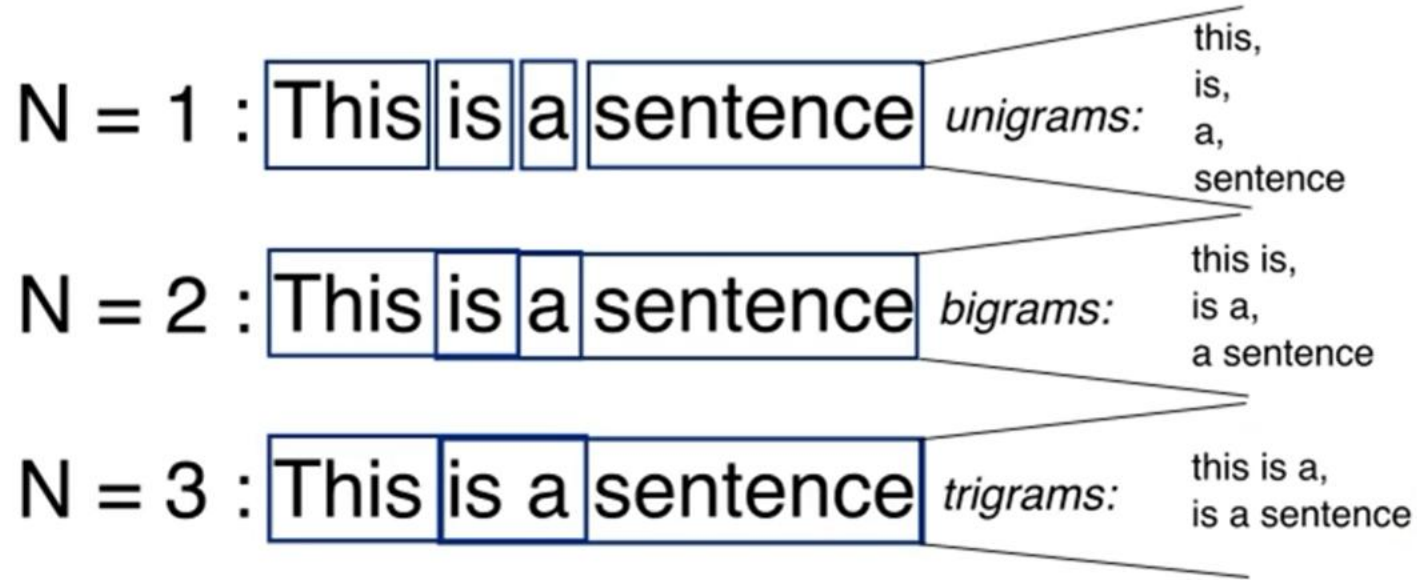
https://alvinntnu.github.io/NTNU_ENC2045_LECTURES/nlp/nlp-pipeline.html

Olasılıksal Dil Modelleri

- N-Gram Modelleri
- Gizli Markov Modelleri (Hidden Markov Models - HMM)
- Maximum Entropy Modelleri (MaxEnt)

N-Gram Modelleri

- N-gram modelleri, bir dizideki (genellikle bir cümledeki) ardışık kelime veya karakter gruplarının olasılıklarını tahmin eder.
- Burada "n", bu gruptaki öge sayısını belirtir.
- Örneğin, bir **unigram** modelinde tek kelimeler, **bigram** modelinde iki kelimelik gruplar, **trigram** modelinde ise üç kelimelik gruplar dikkate alınır.

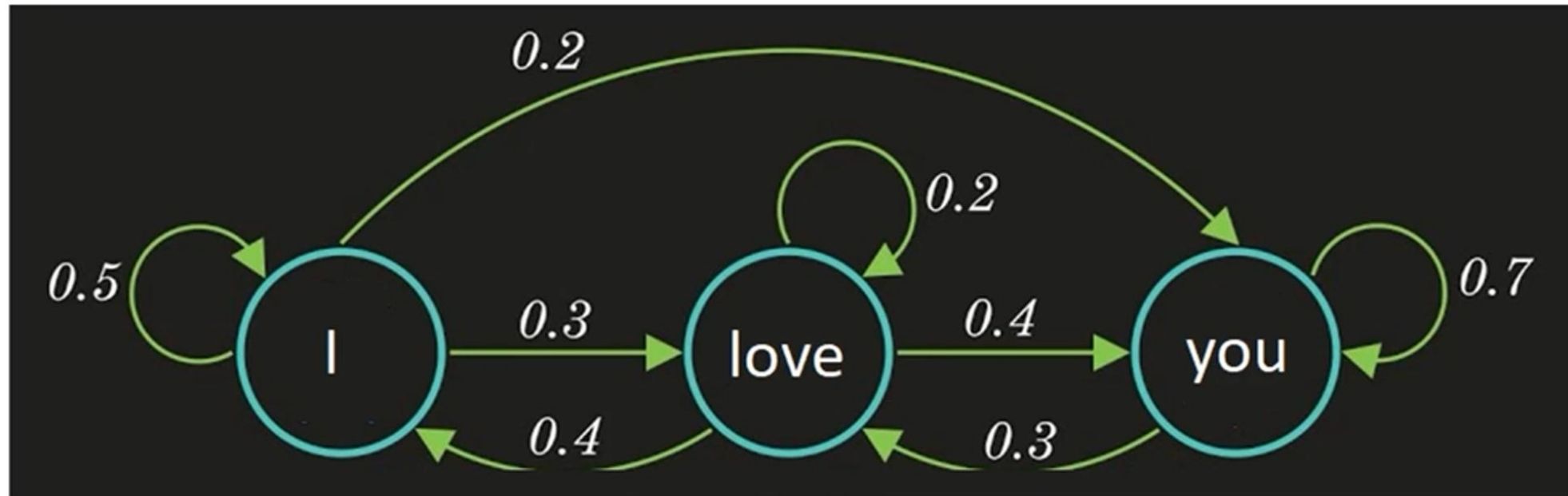


N-Gram Modelleri

- **Kullanım**
 - N-gramlar, bir cümlede bir kelimenin hangi olasılıkla diğer bir kelimedenden sonra geleceğini tahmin etmek için kullanılır.
- **Avantajlar**
 - Basit ve hızlı
 - Yerel bağıntıları iyi yakalar
- **Dezavantajlar**
 - Bağlam Sınırlaması
 - Veri Gereksinimi

Gizli Markov Modelleri (Hidden Markov Models)

- Gizli Markov Modelleri, bir dizi gözlemin (kelimeler veya karakterler gibi) arkasında gizli bir durum dizisinin (örneğin, dil bilgisel kategoriler) olduğu varsayımına dayanır.
- HMM, her bir durumun belirli bir olasılıkla başka bir duruma geçeceğini ve her durumun belirli bir gözlemi üreteceğini varsayar.

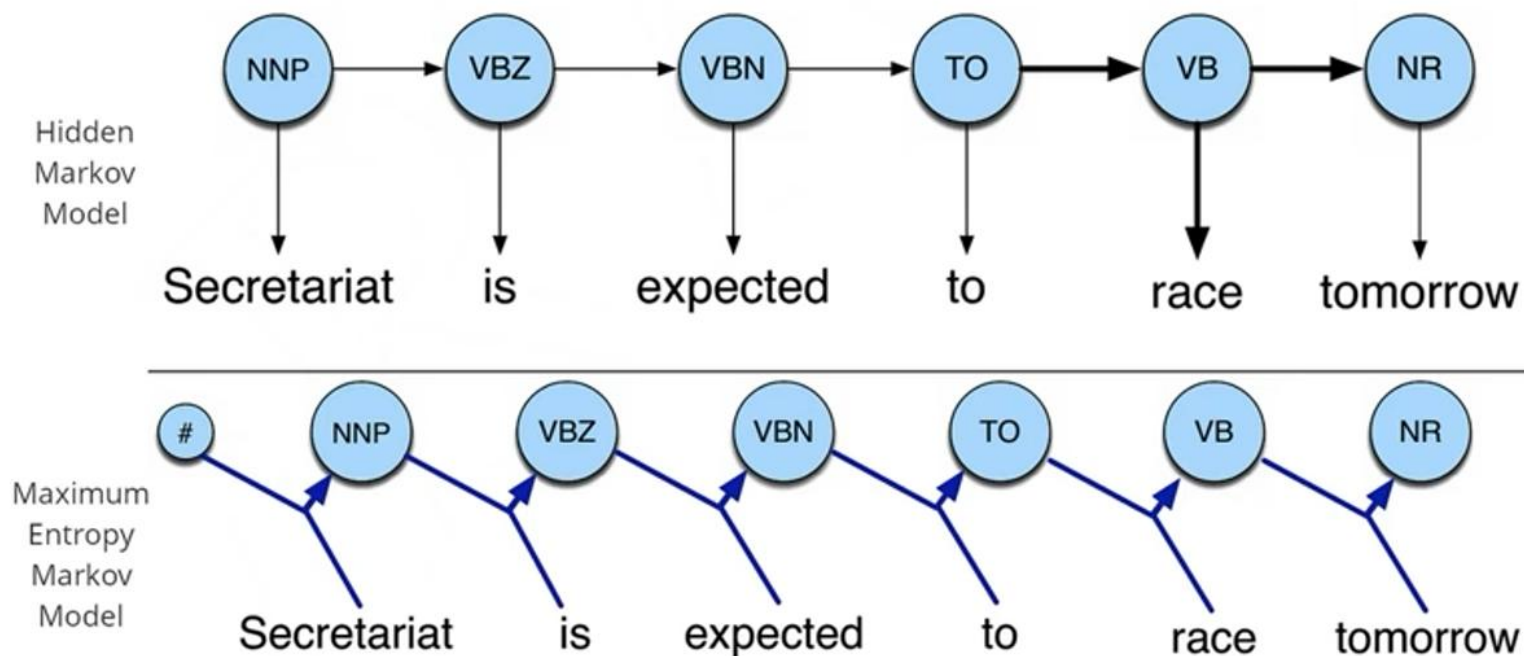


Gizli Markov Modelleri (Hidden Markov Models)

- **Kullanım**
 - Konuşma tanıma
 - Dil modelleme
 - Parça etiketleme (Part-of-Speech Tagging)
- **Avantajlar**
 - Bağlam Modelleme
 - Verimli Algoritmalar
- **Dezavantajlar**
 - Basitleştirici Varsayımlar
 - Eğitim Zorluğu

Maximum Entropy Modelleri (MaxEnt)

- Maximum Entropy (Maksimum Entropi) Modelleri, bir olasılık dağılımı tahmin ederken mümkün olduğunca az varsayımda bulunmayı hedefler.



Maximum Entropy Modelleri (MaxEnt)

- **Kullanım**
 - MaxEnt modelleri, özellikle sınıflandırma görevlerinde kullanılır.
 - Örneğin, bir cümlenin belirli bir sınıfa (pozitif, negatif duygu gibi) ait olma olasılığını tahmin edebilir.
- **Avantajlar**
 - Esneklik
 - İyi Genelleme
- **Dezavantajlar**
 - Hesaplama Maliyetleri
 - Özellik Mühendisliği