



ترجمه ماشینی

کارگاه یادگیری عمیق با پایتون

سید ناصر رضوی www.snrazavi.ir

۱۳۹۷

پردازش زبان طبیعی

> do not tell me what to do
= به من نگو چه کار کنم
< به من نگو چه کار کنم

> the government banned cigarette advertising on television
= دولت تبلیغات سیگار در تلویزیون را ممنوع کرد
< دولت تبلیغات سیگار در تلویزیون را ممنوع کرد

> i really want to be with tom right now
= واقعا دلم می خواهد هم اکنون با تام باشم
< واقعا دلم می خواهد هم اکنون با تام باشم

> let it go
= ولش کن
< ولش کن

> tom asked several people the same question
= تام از چندین نفر، سوال مشابهی را پرسید
< EOS> تام از چندین نفر، سوال سوال را پرسید

□ دسته بندی متون

□ مدل سازی زبان

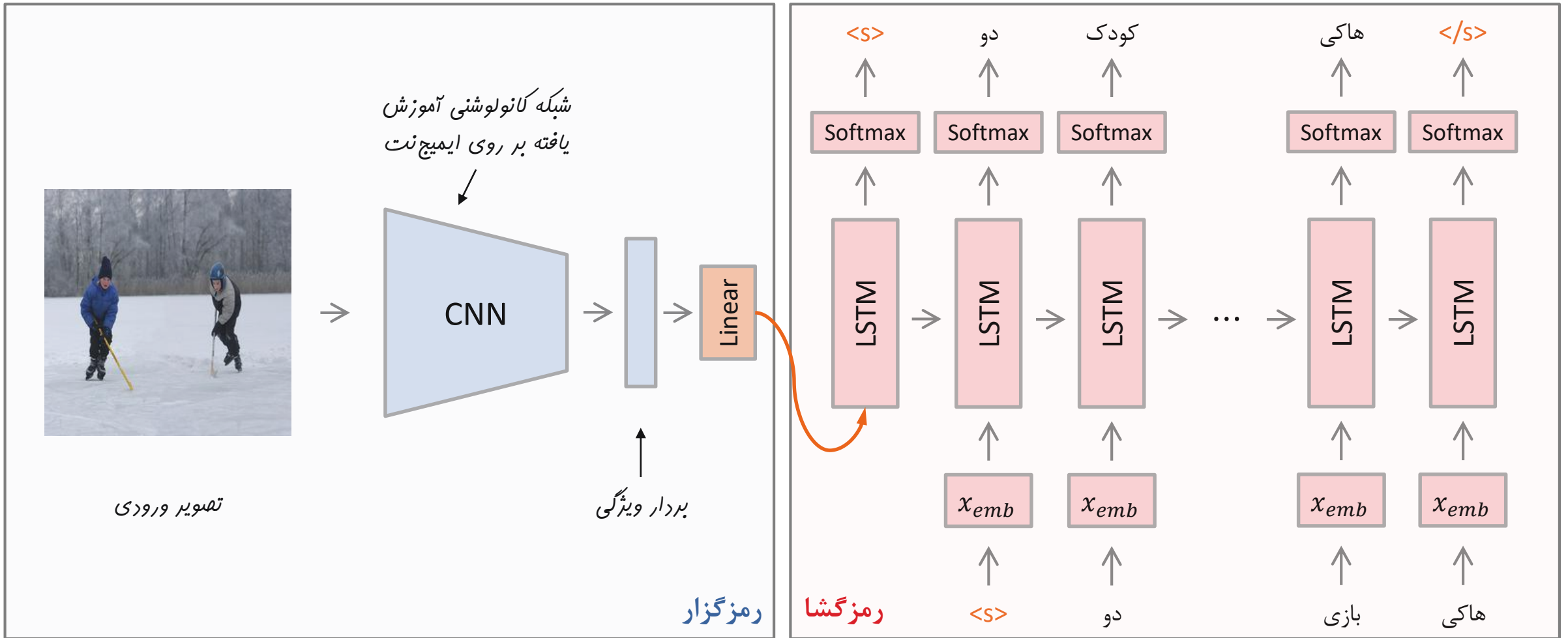
□ عنوان بندی تصاویر

□ ترجمه ماشینی

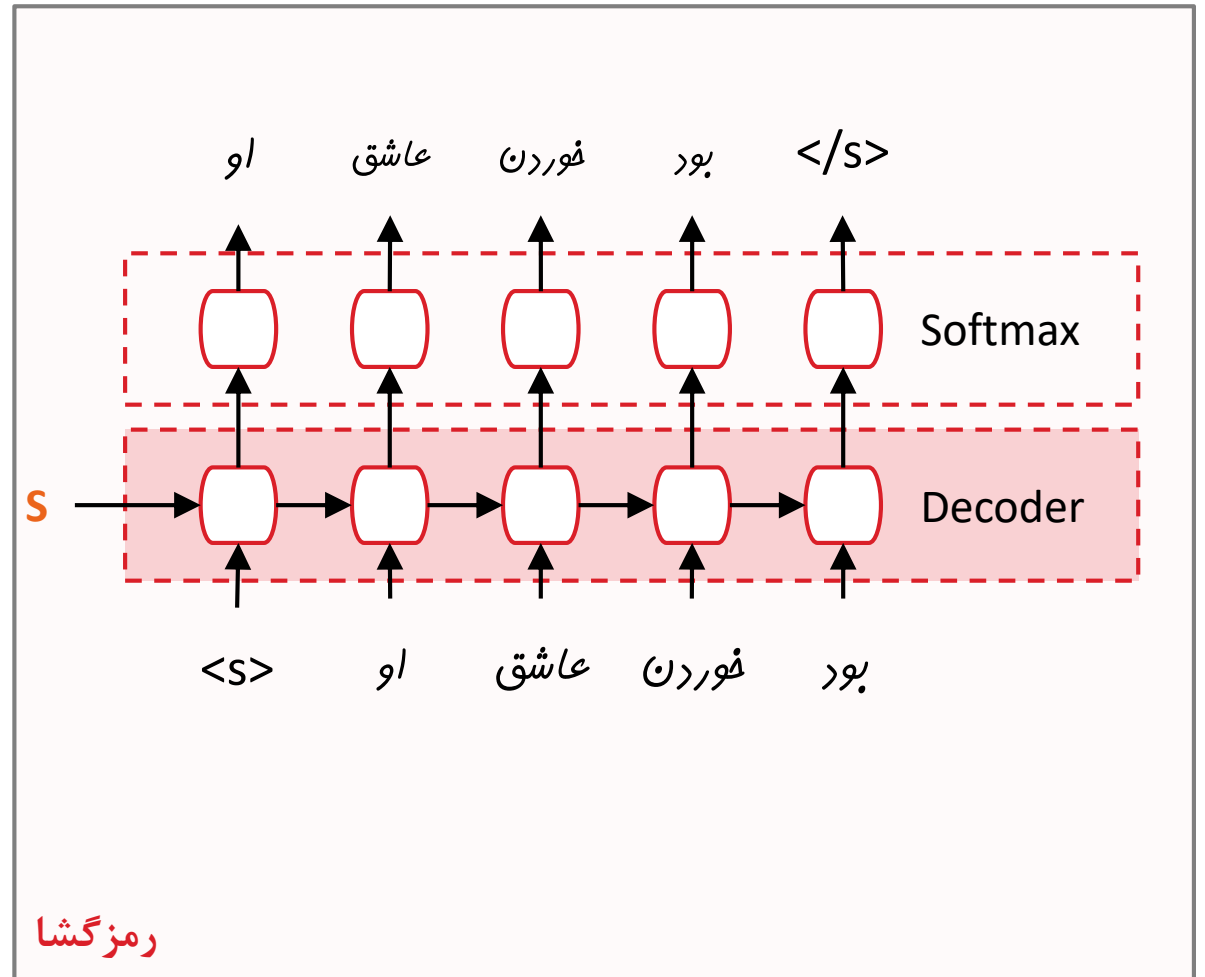
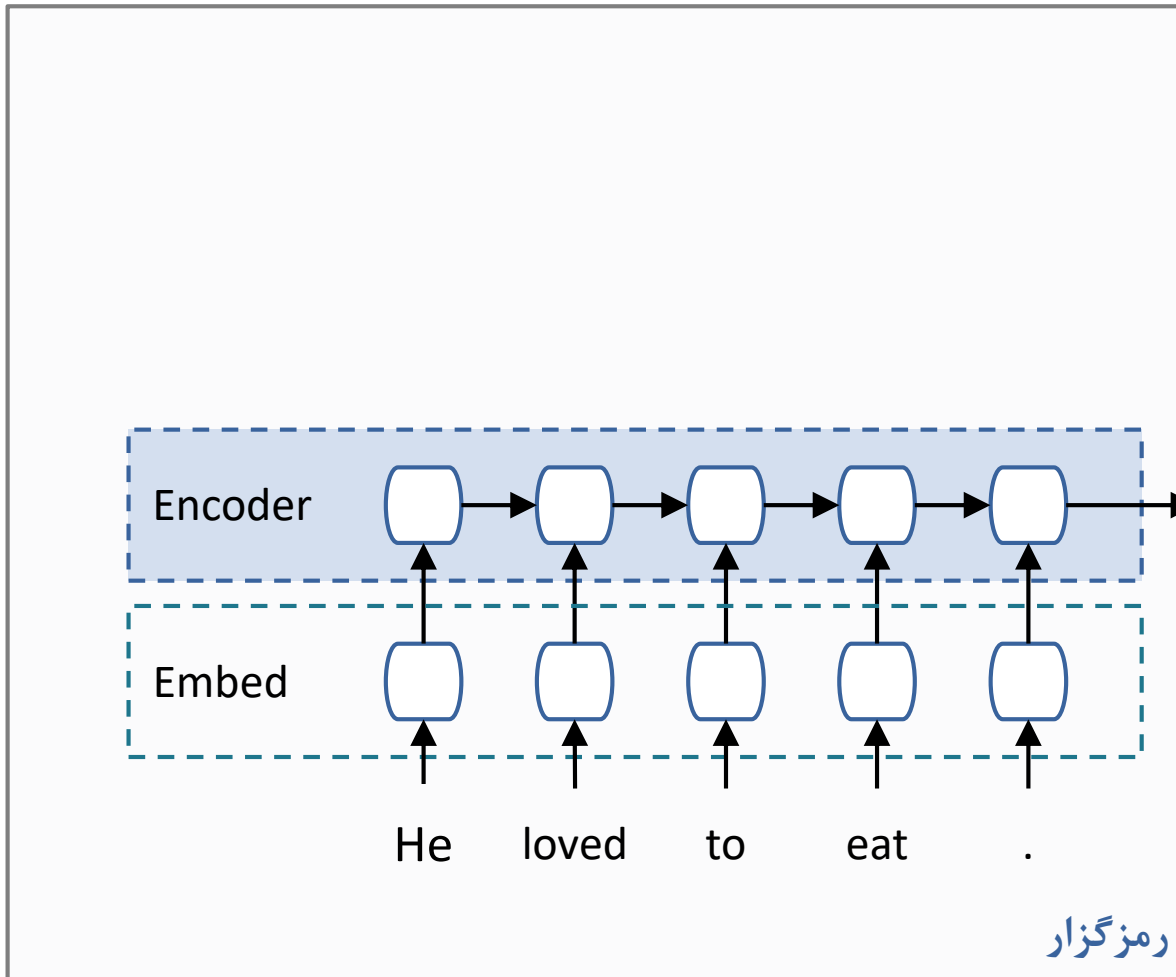
فهرست مطالب

- ترجمه ماشینی
- دادگان: پیکره موازی
- معماری دنباله به دنباله (رمزگزار-رمزگشا)
- بهبود نتایج
- سازوکار توجه
- اجبار یادگیری

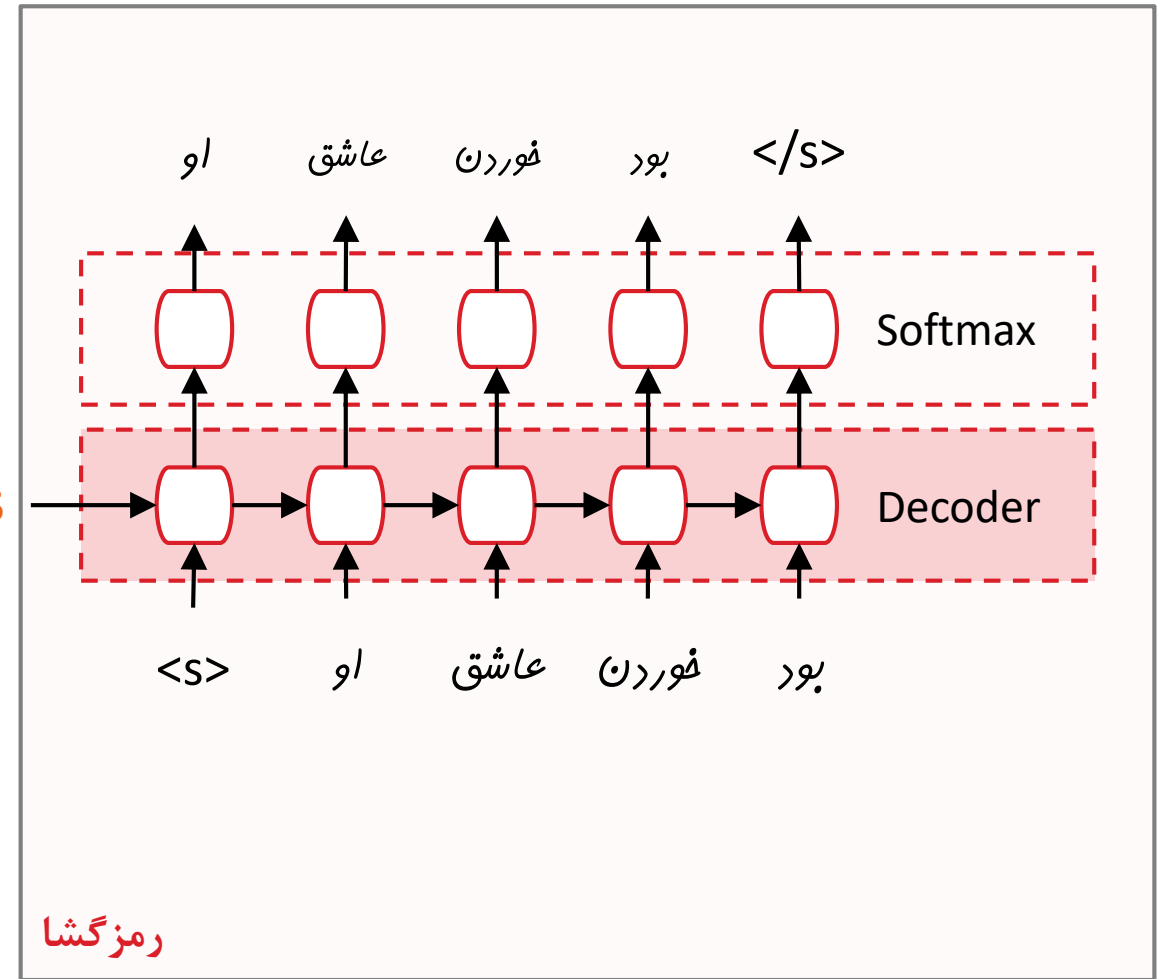
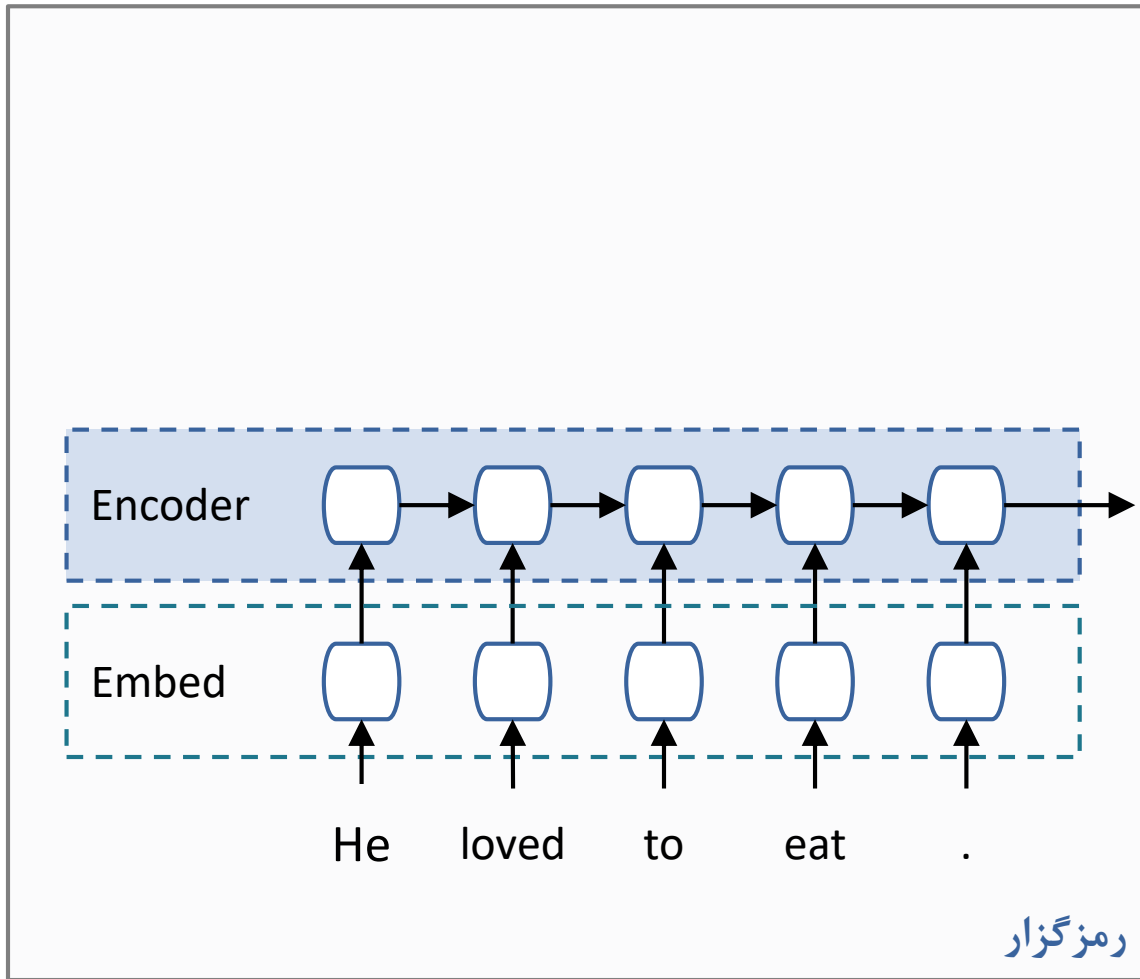
یادآوری: معماری رمزگزار-رمزگشا



معماری دنباله به دنباله

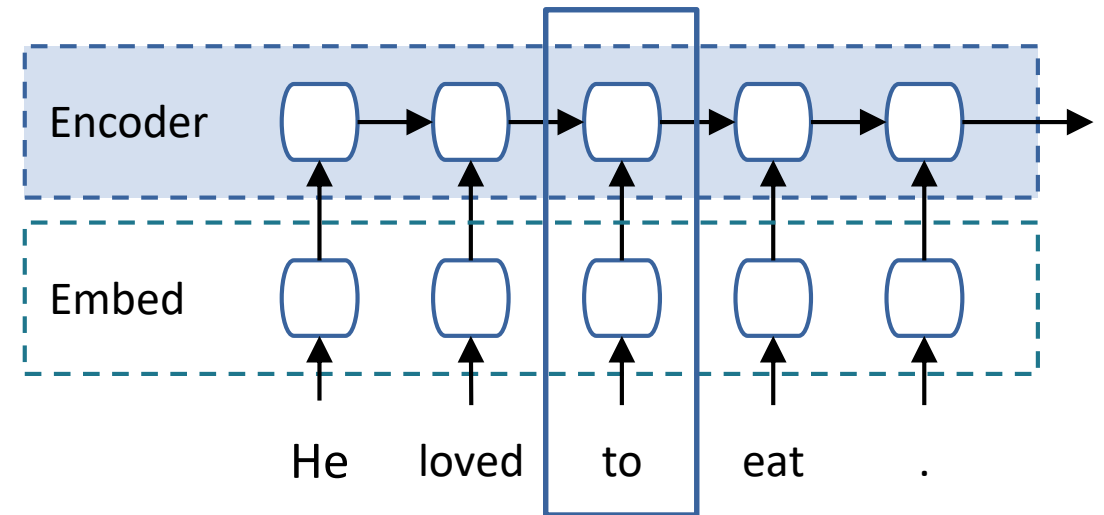


معماری دنباله به دنباله

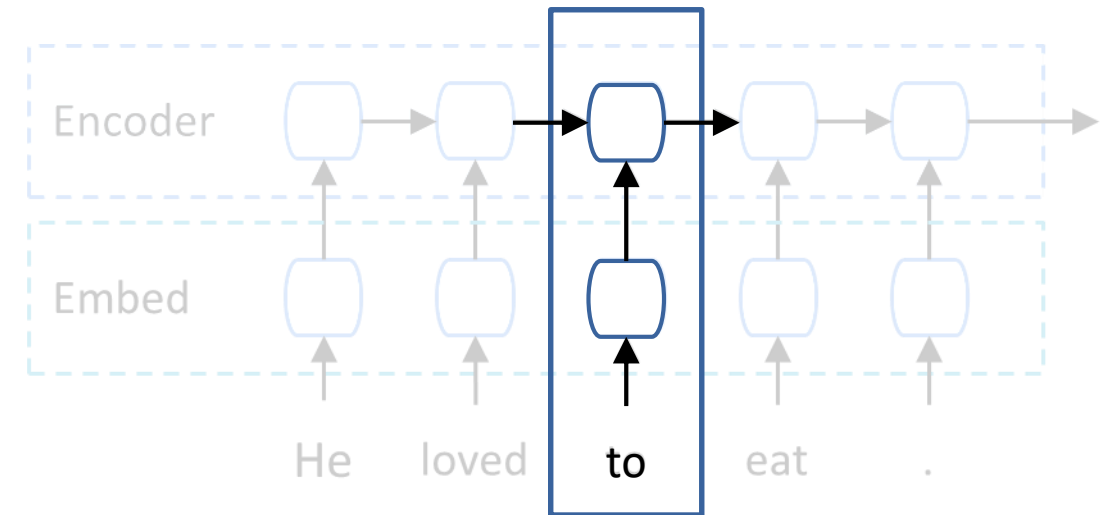


```
def forward(self, x, h):  
    x = self.embedding(x)  
    x = self.dropout(x)  
    x, h = self.gru(x, h)  
    return x, h
```

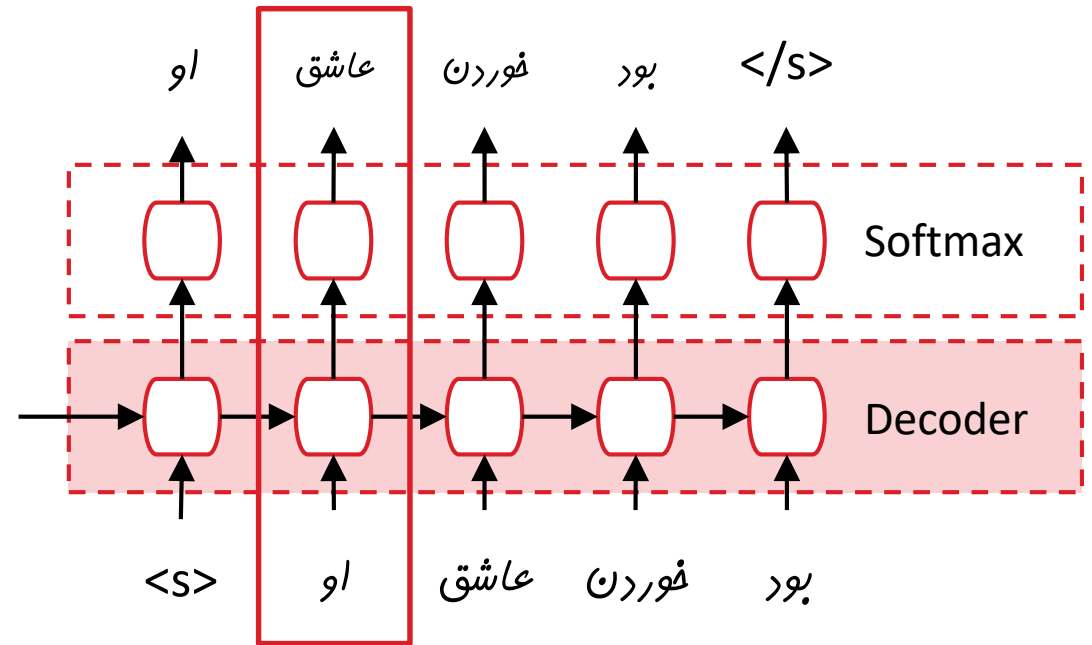
دوسویه و دارای عمق دلخواه



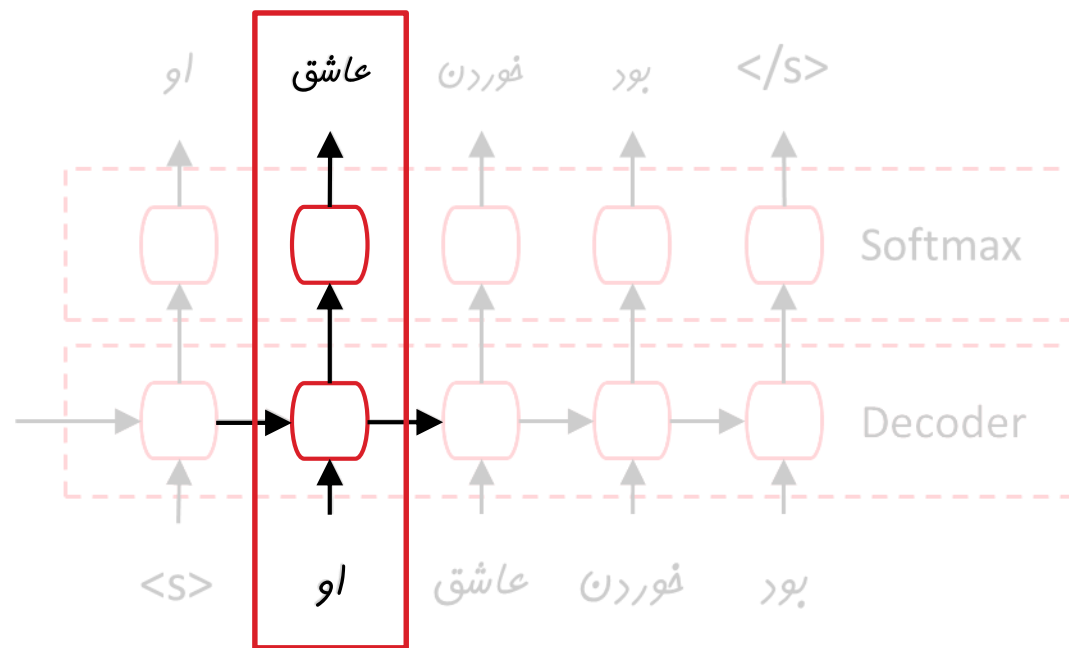
```
def forward(self, x, h):  
    x = self.embedding(x)  
    x = self.dropout(x)  
    x, h = self.gru(x, h)  
    return x, h
```



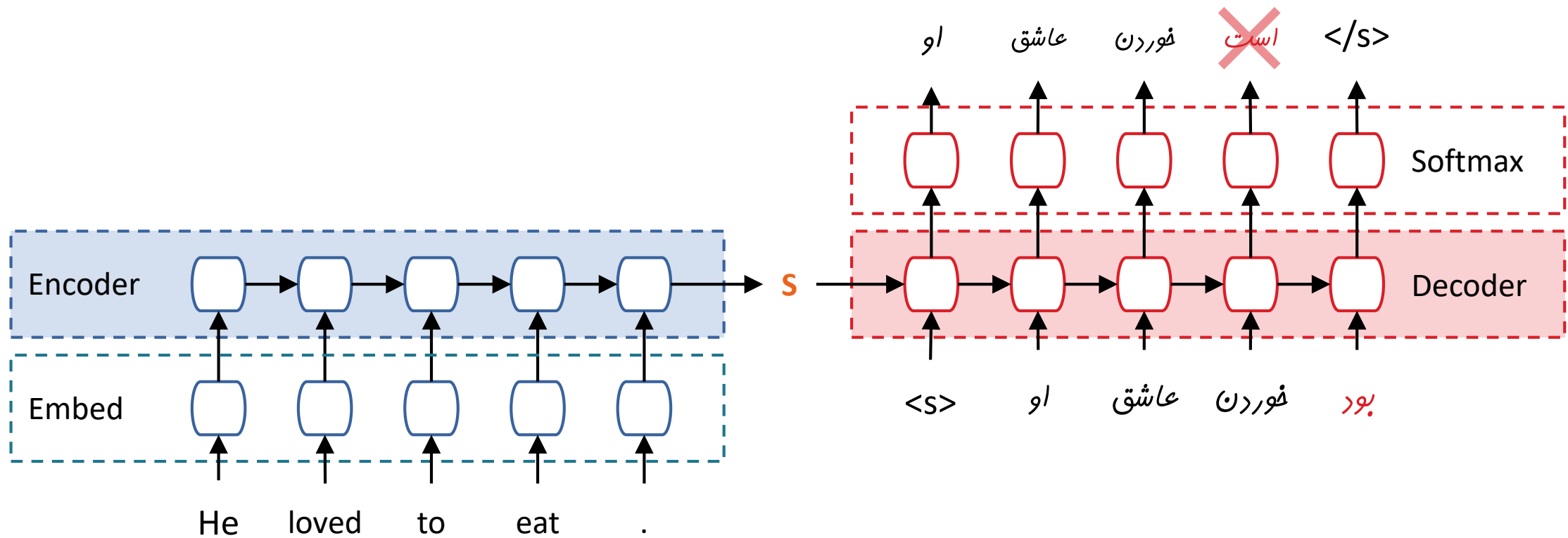

```
def forward(self, x, h):
    x = self.embedding(x)
    x = self.dropout(x)
    x, h = self.gru(x, h)
    x = self.dropout(x)
    x = self.fc(x)
    return x
```



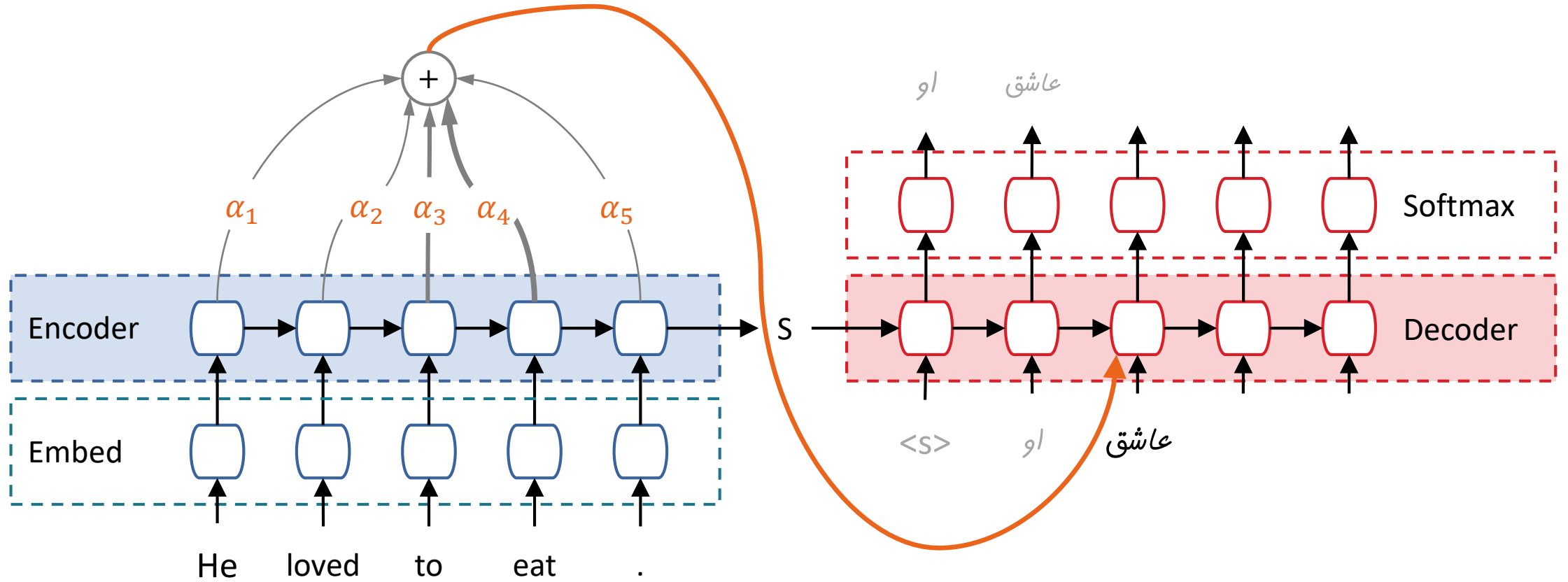
```
def forward(self, x, h):
    x = self.embedding(x)
    x = self.dropout(x)
    x, h = self.gru(x, h)
    x = self.dropout(x)
    x = self.fc(x)
    return x
```



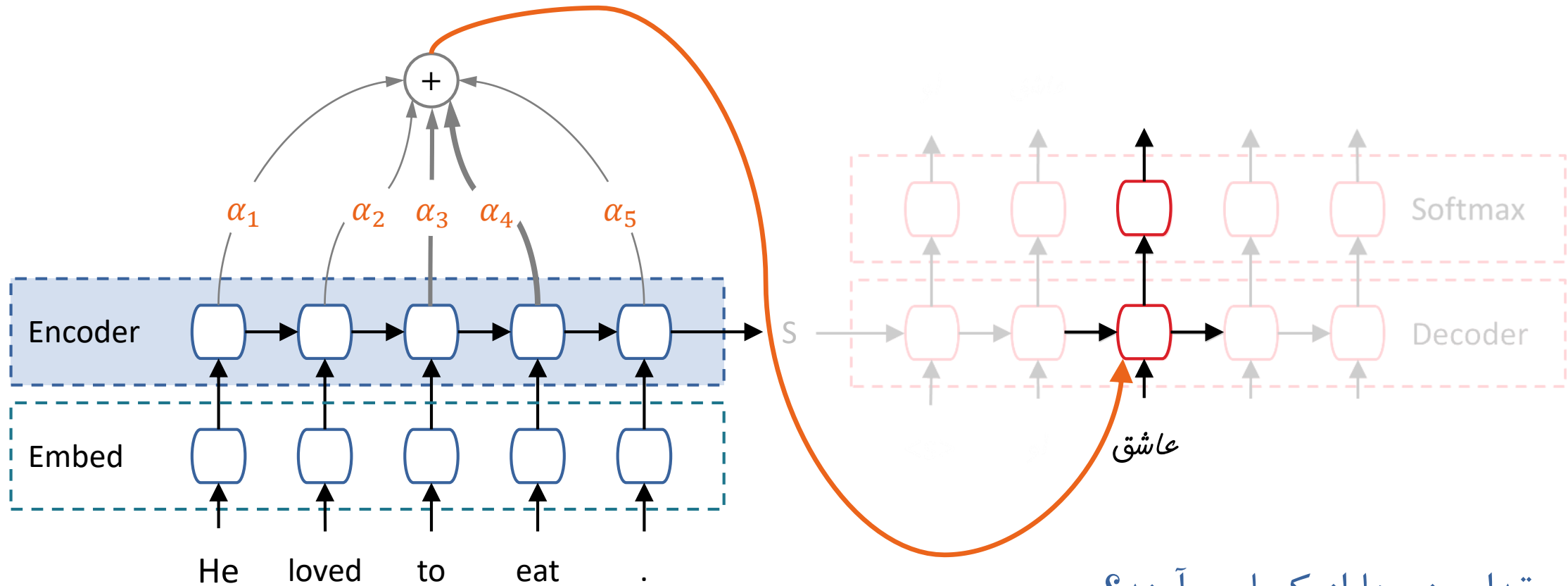
اجبار یادگیری



ساز و کار توجه



ساز و کار توجه



مقدار وزن ها از کجا می آیند؟



- استفاده از پیکره موازی با جملات بیشتر
- استفاده از بردارهای واژگان اولیه
- استفاده از جستجوی دسته‌ای برای تولید جملات
- افزایش عمق و اندازه لایه‌های برگشتی

چالش‌ها

- تعداد واژگان بسیار زیاد
- چگونگی برخورد با کلمات نادر
- راه‌حل. مدل‌سازی در سطح کاراکتر و زیرکلمه