在数说故事的实习，应该说收获最大的是寻找资料的方法，我发现有两个超级神器，1）谷歌，在谷歌学术了可以查阅到非常多前沿的论文，阅读论文可以为解决问题提供思路，这个非常好；2）GitHub，这个神器可以查阅到很多模型的开源代码，以前觉得做研究什么东西都需要从头开始手动写，其实有很多开源的代码提供使用的；

应该说，我在数说做的两个月，一般的过程是：首先分析问题，看看当前的业务可以使用哪种类型的方法解决；然后做一下论文的review，看看别人一般是用什么方法来，解决这些问题的；接着精细地寻找想用来实现的论文，并且看看是否能在GitHub上找到开源的代码。

* **词向量：**

词向量的开山鼻祖

A Neural Probabilistic Language Model

http://www.jmlr.org/papers/volume3/bengio03a/bengio03a.pdf

2013年提出，也是现在最常用的w2v

#论文写得很垃圾，还是看博客比较好！

Distributed Representations of Words and Phrases and their Compositionality

Efficient Estimation of Word Representations in Vector Space

2014年的词向量方法

GloVe: Global Vectors for Word Representation

<https://nlp.stanford.edu/pubs/glove.pdf> #论文下载地址

https://github.com/stanfordnlp/GloVe#download-pre-trained-word-vectors

#预训练的glove词向量下载地址（一般都是直接使用预训练的）

腾讯那篇：

Directional Skip-Gram: Explicitly Distinguishing Left and Right Context for Word Embeddings

http://aclweb.org/anthology/N18-2028

Fasttext也可以做词向量，要找找那篇论文看看：

2018年，动态词向量ELMO（这个看论文用得蛮多的）

Deep contextualized word representation

采用Fine-Tuning方式的GPT

业界非常认同的bert:

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

https://arxiv.org/pdf/1810.04805.pdf

* **序列标注：**

Neural Architectures for Named Entity Recognition（bilstm-crf）

https://arxiv.org/pdf/1603.01360.pdf

* **aspect level sentiment analysis：**

Deep Learning for Sentiment Analysis

https://arxiv.org/pdf/1801.07883

Recent trends in deep learning based natural language processing

https://arxiv.org/pdf/1708.02709

Multi-grained Attention Network for Aspect-Level Sentiment Classification

aclweb.org/anthology/D18-1380

Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks

<https://arxiv.org/pdf/1804.06536.pdf>

Interactive Attention Networks for Aspect-Level Sentiment Classification

https://arxiv.org/pdf/1709.00893.pdf

Recurrent Attention Network on Memory for Aspect Sentiment Analysis

<http://www.aclweb.org/anthology/D17-1047>

Aspect Level Sentiment Classification with Deep Memory Network

https://arxiv.org/pdf/1605.08900.pdf

Transformation Networks for Target-Oriented Sentiment Classification

<https://arxiv.org/pdf/1805.01086.pdf>

Attention-based LSTM for Aspect-level Sentiment Classification

Content Atention Model for Aspect Based Sentiment Analysis

IARM\_Inter-Aspect Relation Modeling with Memory Networks in Aspect-Based Sentiment Analysis

* **Unsupervised Aspect Extraction：**

An Unsupervised Neural Attention Model for Aspect Extraction

http://aclweb.org/anthology/P/P17/P17-1036.pdf

* **文本分类：**

Hierarchical Attention Networks for Document Classification

https://www.cs.cmu.edu/~diyiy/docs/naacl16.pdf

* **attention：**

Attention Is All You Need

https://arxiv.org/pdf/1706.03762.pdf

* **超级无敌XGBOOST**

XGBoost: A Scalable Tree Boosting System

https://www.kdd.org/kdd2016/papers/files/rfp0697-chenAemb.pdf