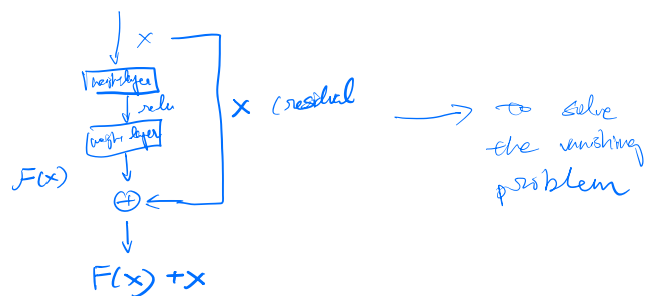


- Solve for this
- deep convnet \rightarrow gradient vanishing when back prop
 - Residual Networks



Natural Language Processing :

Global Word representation :

- Word Net
 - one-hot vectors
- vector size fixed
sparsity
no natural notion of similarity.

word-embedding matrix?

PPMI matrix \rightarrow very sparse.

for sparsity : do SVD

minimize $\| M^{PPMI} - W V^T \|_F^2$

$W: V$ $W \times d$ $W \times d$ $d \times d$

for binous
elium wife summer

\rightarrow redundant info within the sparse matrix.

SVD: compression

word2vec

on page 25
difference between two group of words,
if similar relations,
vector looks similar

problem: Bank of China v. river bank

\rightarrow solution: contextual word embedding.

mission: text classification.

ARV \rightarrow simplified
of
LSTM