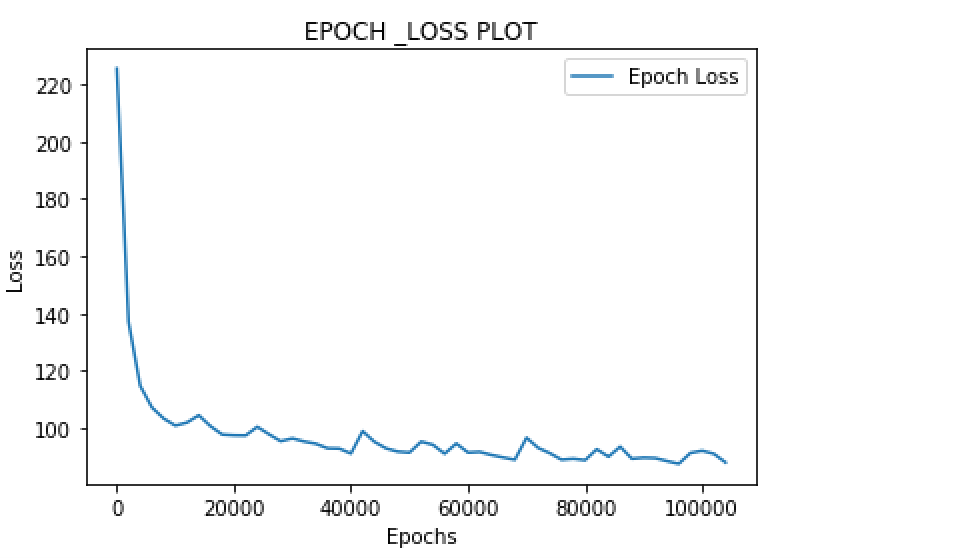
**Task 1**

**Training loss vs epochs plot**



**Output of the network**

veres

profer cosker, and in the

veresta, ime eve cowh,” am

foratiend of the Rurra’s beakryey mur sorvame ward was cowtwe

cove beeste consund,” prof cabit coice of thinain tir whe he sopctout to the the grunet ther had percarntnes wivaexcove tome cas lon to diantye but hout, wher’sswor plas ho ques a

plon. A bat ancspemes

reapen

sos.

“Mishan.

“Abutt apisut wipfrin, In!” recterabn muth son ricle pick of thipe eat sen ofly

son

that eve oficksing stade as thealicene surlor, dea mimt exutur exshevek fomentes

whe anamgly theat naes ant wats bednd it thew bald sian, chew the kpey thad womere thotercrectearq\_ fade he retwe isaadpl anghe that, tho oand the this Zfofart, mlencit of the gould to thee comteming tI that dablics the inge-to vich Jo In his us protutcowit, of the abe Bive ede.”

bat nod durd

mid tuska ad oon abs, war.

OPhys the sadis saaly.

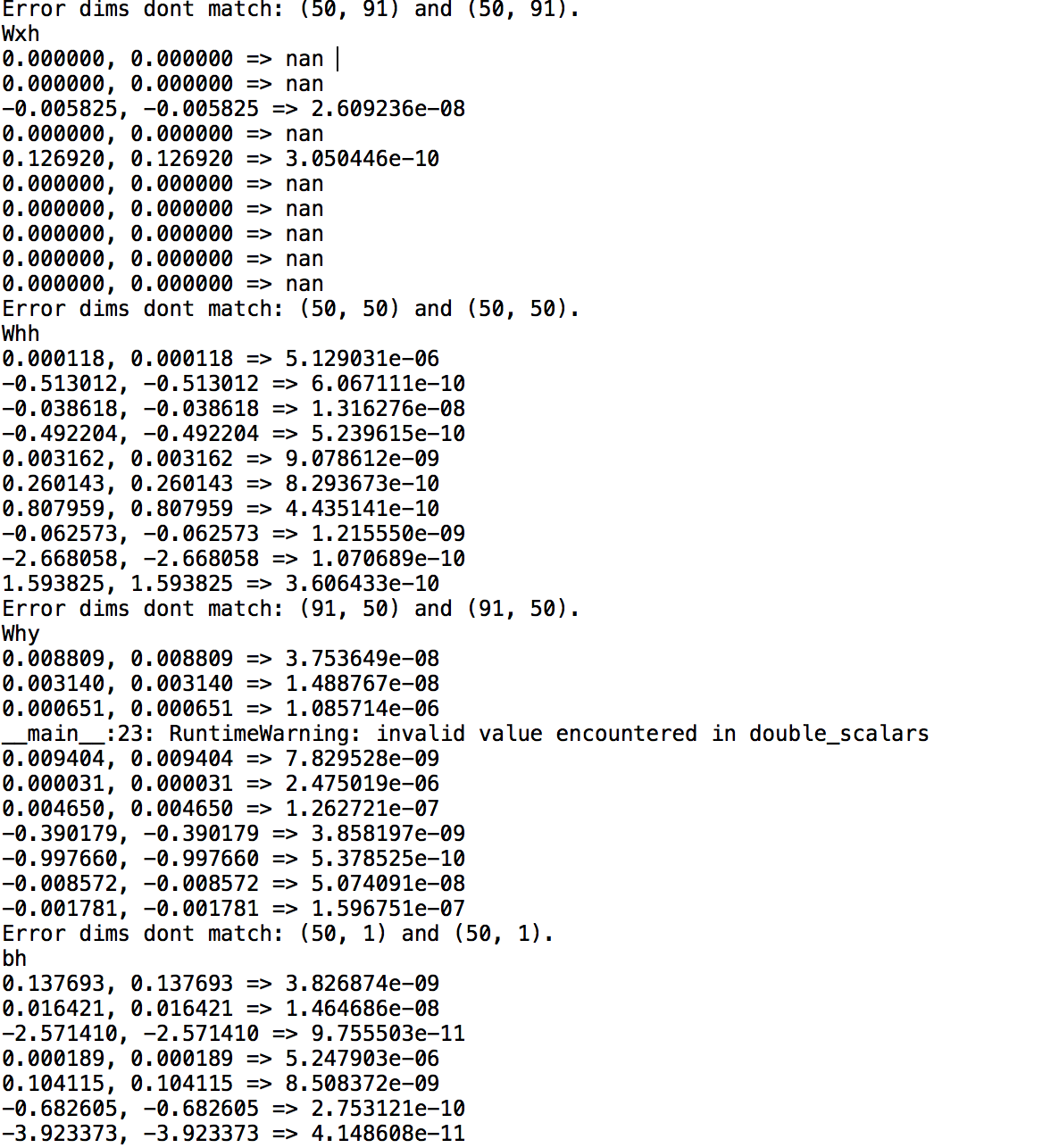
The behy to tood op fat bugu to colterun se expedtemry che the tarisinove horficasly.”

“lo the “Mestory wat tiry -Orpeat the erpocle; “Whe free har faln sour, frelrolled torscovyt. Anssres in offale urespalt wotedstery thached rexinster a mest!” sind bomemstimm!” lroumvanl to themrette povensiaw sa!” rices

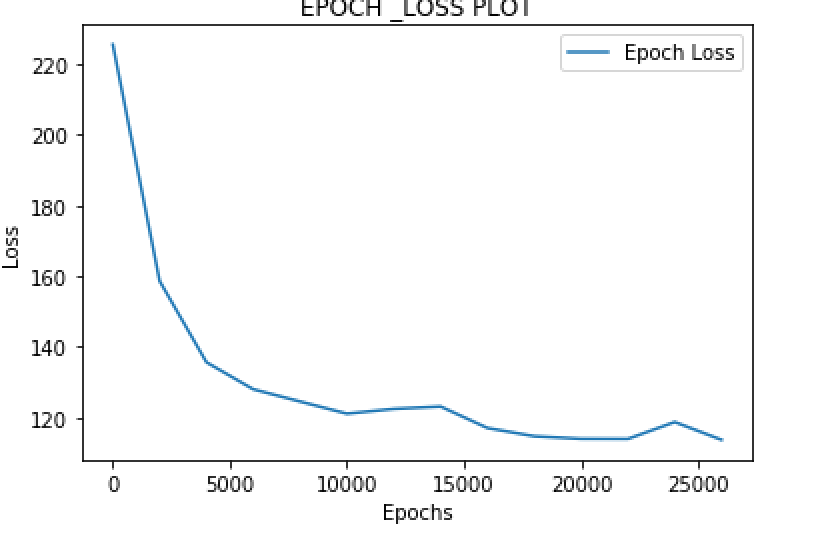
llines, the

reavofssey the vavi. ?”

**Gradient Check Routine**

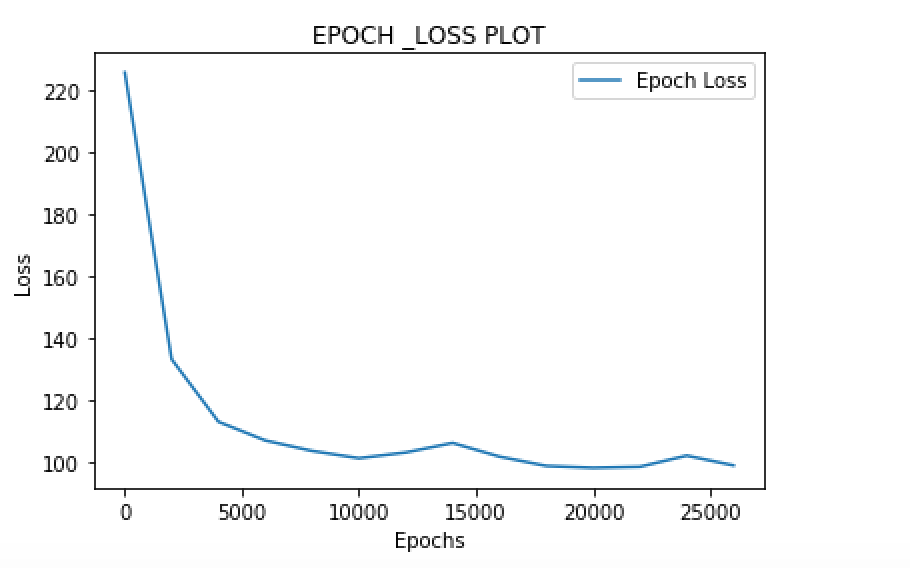
****

**Doubling the number of hidden units**

****

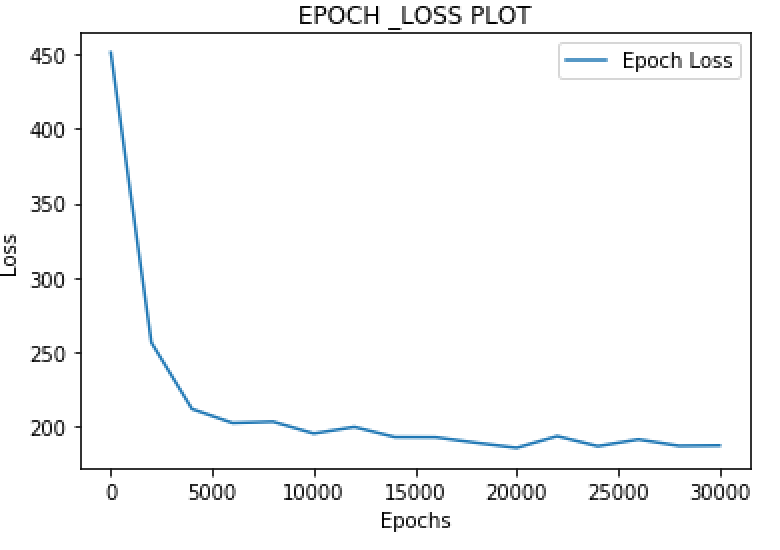
By examining the graph and the output text loss is gradually decreased with the epochs and the prediction text output has some meaning.

**Halving the number of hidden units**

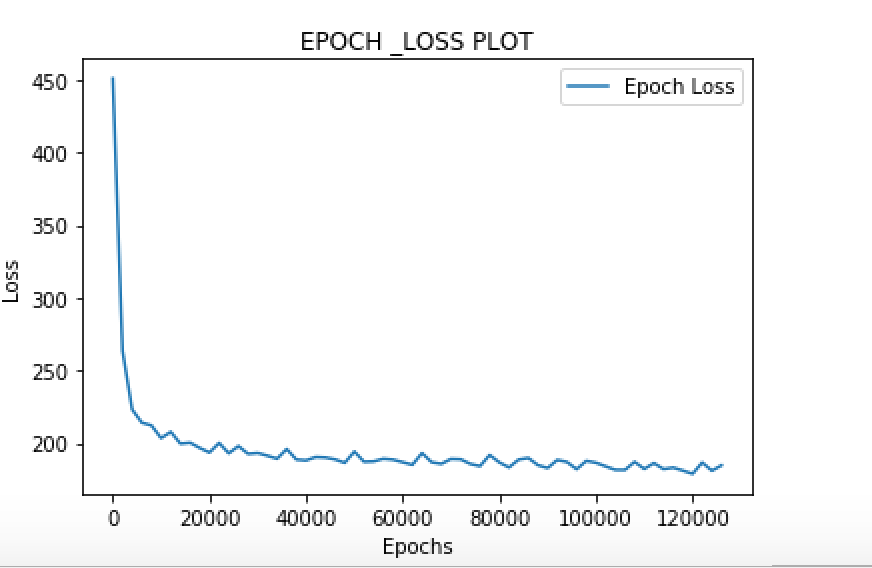
****

The text output is random and has less meaning and the loss is decreased in the same way as the doubling the hidden units

**Doubling the sequence length**

****

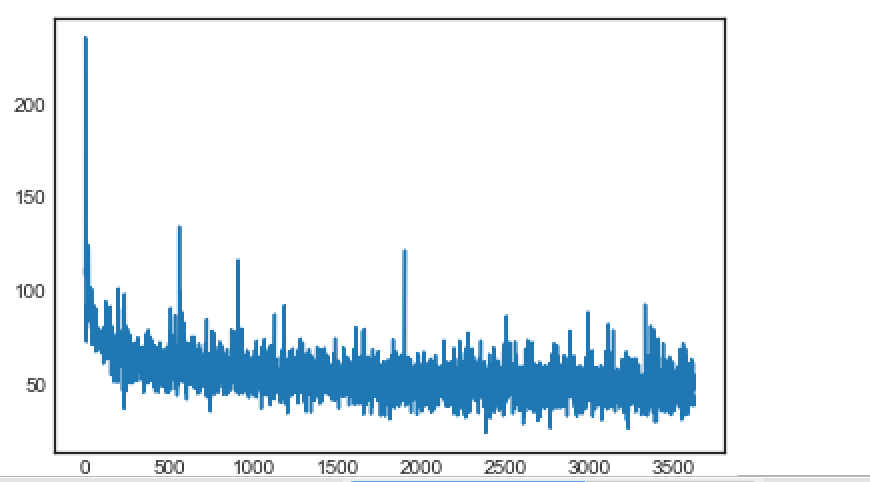
Initially loss is more and is gradually decreased with the epochs and the predicted text output don’t have any proper meaning when the sequence length is doubled

**Halving the sequence length**

Initially the loss is more and is decreased and the loss has not much difference and the predicted text output from the iterations has some meaning.

**Task 2**

**Training loss vs epochs plot**

****

**Output of the network**

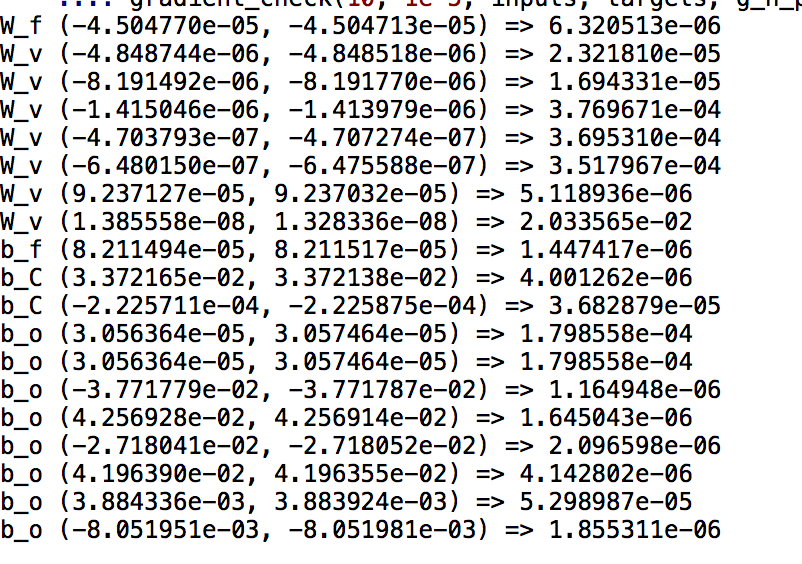
runtjut is thany caruic sime, the visteots

abd a wer in the

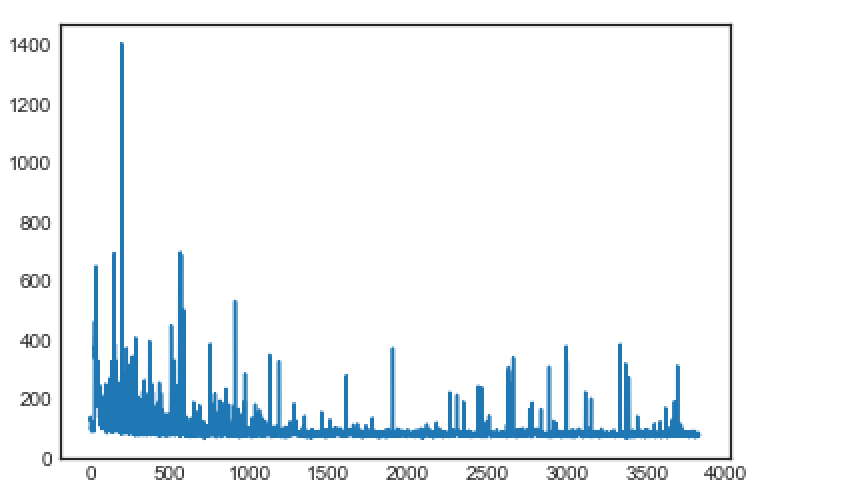
towthal the you seut it lefcions lescomperon Servadat s2ie consered veced the foubl muc, bur oo has the hitt.

0evooun was liognoin, the cu

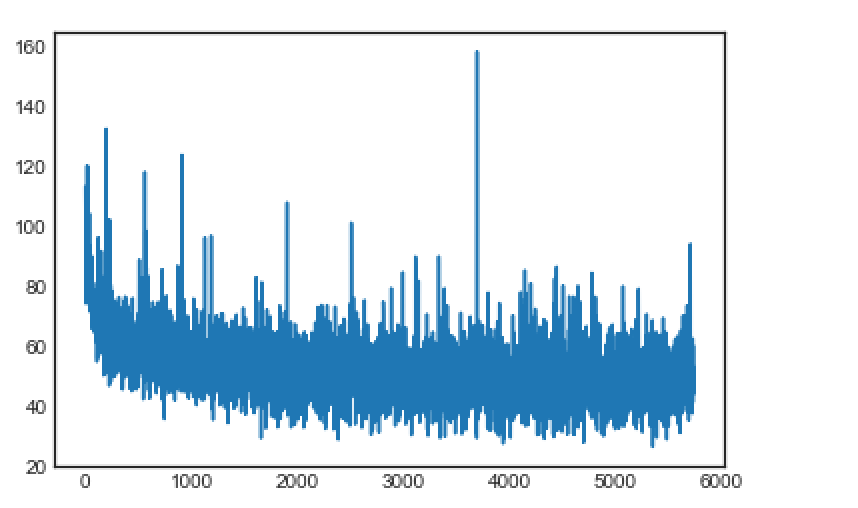
**Gradient Check Routine**



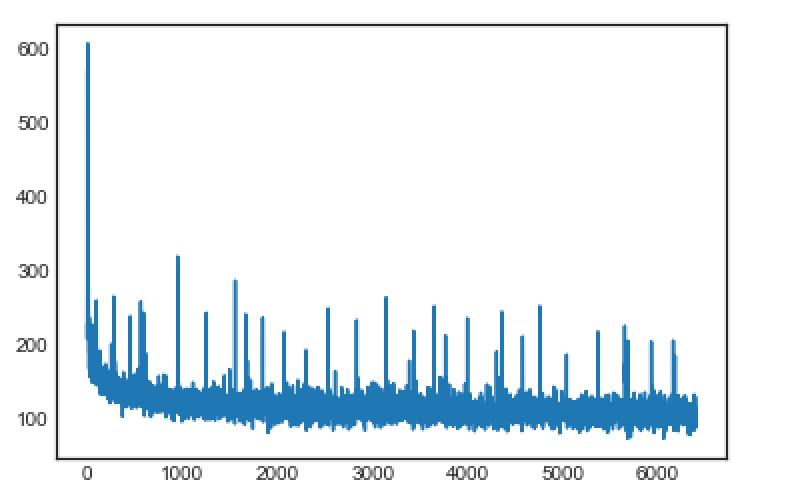
**Doubling the number of hidden units**



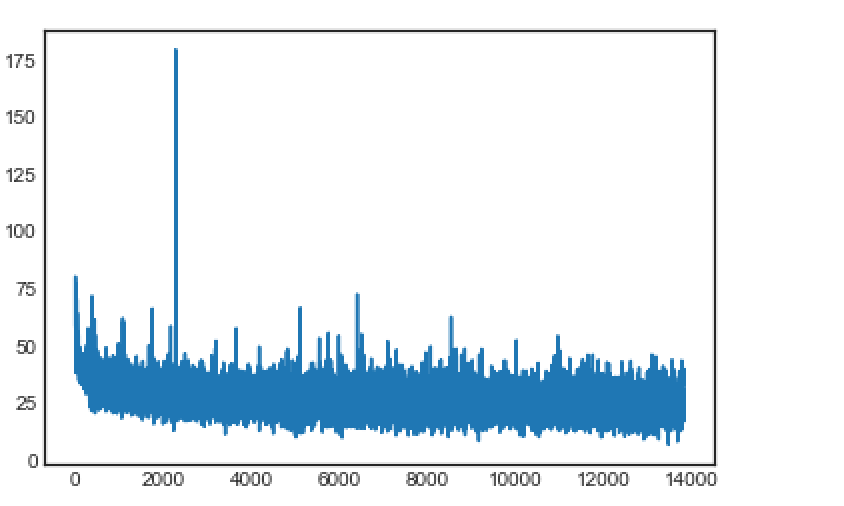
**Halving the number of hidden units**



**Doubling the sequence length**



**Halving the sequence length**

****