Web Mining Additional Learning

## Name: Om Ashish Mishra

## Registration Number: 16BCE0789

## Slot: F2

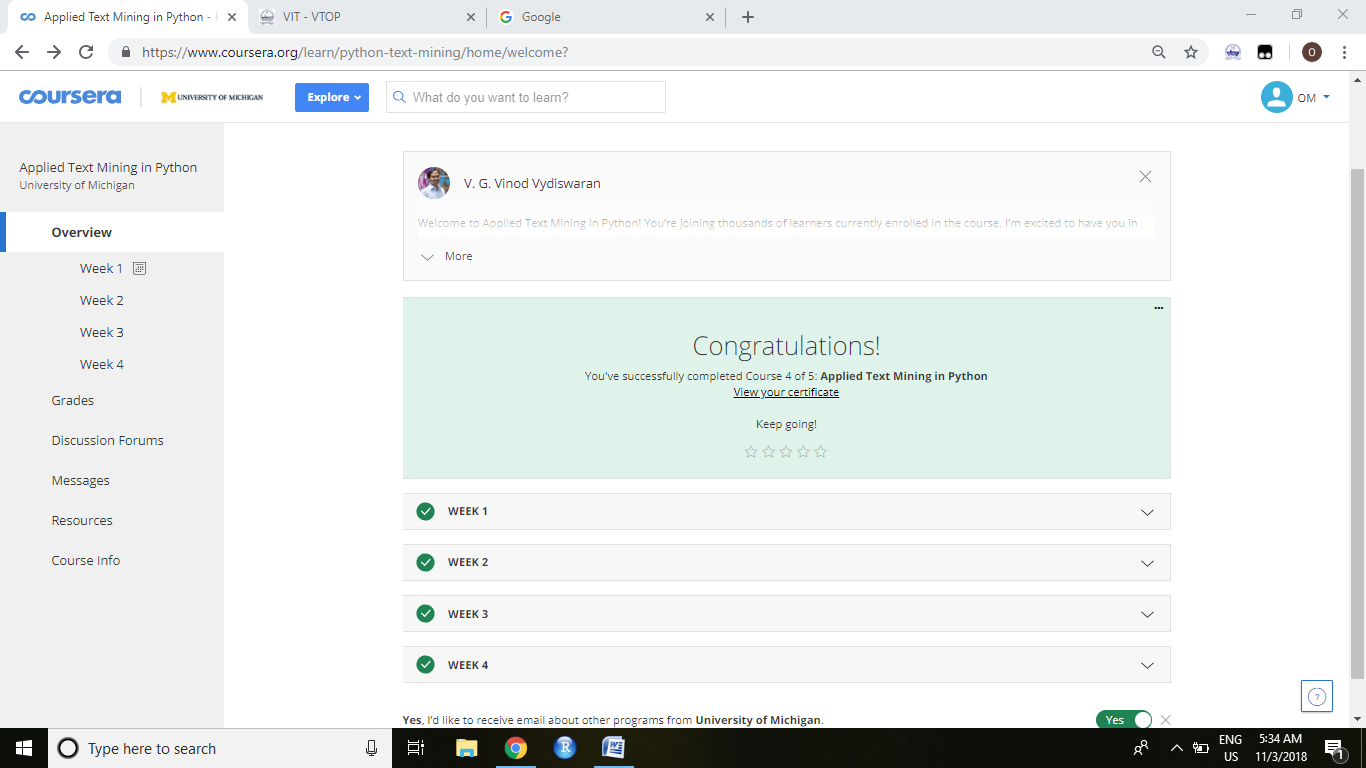
## Submitted to Professor: Lokesh Kumar R

## Applied Text Mining in Python

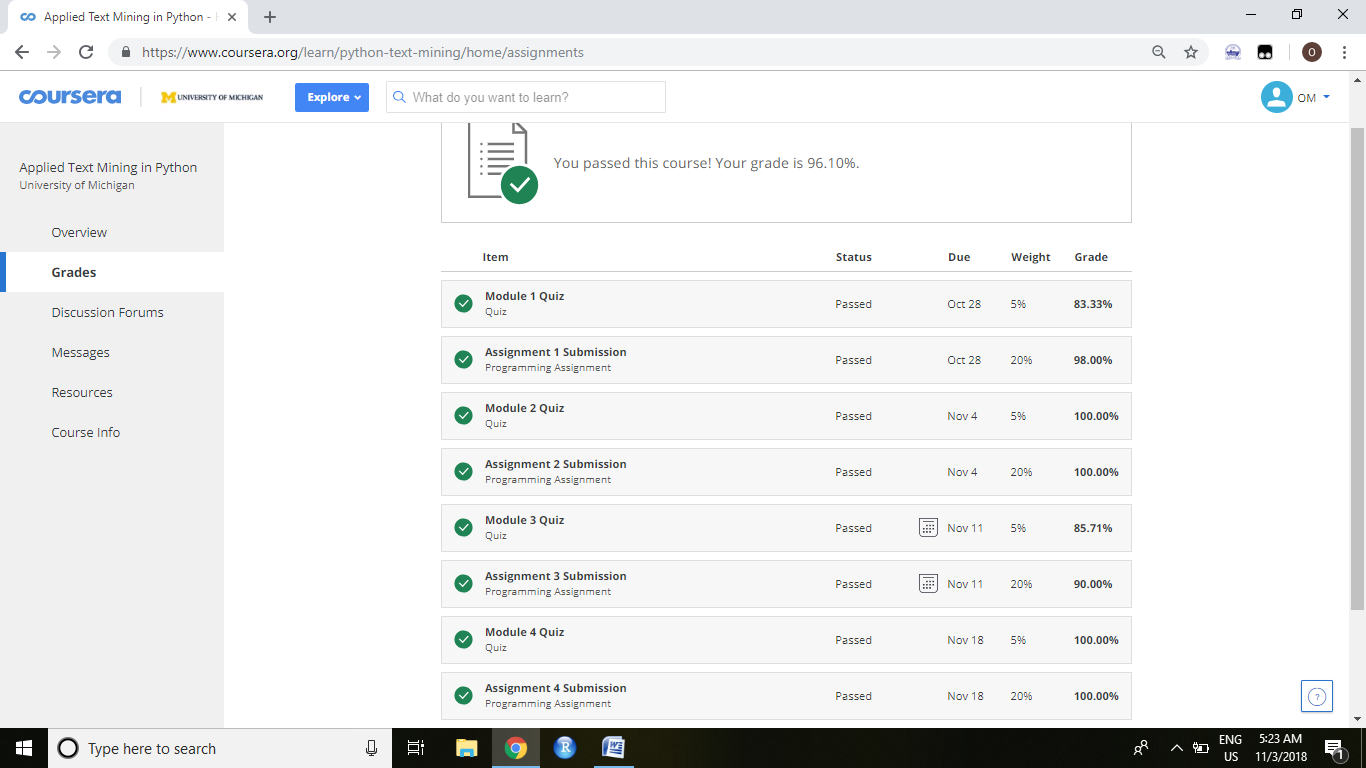
<https://www.coursera.org/account/accomplishments/certificate/5K62SMGKGZSE>



# Course Completion of the 4 Weeks:-



# Grade Completion by qualifying the Assignments and Quizzes:-



# What is learnt week wise distribution:-

**Week – 1**

This week includes -

Interpret text in terms of basic building Blocks : Sentences and words.

Identifying common problems with raw text and perform text cleaning tasks in python.

Write regular expressions to find textual patterns.

**Week – 2**

This week includes -

Describe different natural language tasks.

Process free text through the NLTK toolkit to tag language constructs onto text.

Derive meaningful features from text.

**Week - 3**

This Week includes –

Compare text classification to other classification approaches (covered in Applied

Machine Learning in Python as well)

Describe the Naive Bayes and Support Vector Machine algorithms

Classify text in two classes using one of these approaches in Python

Identify and extract features from text and transform them into feature vectors for

the machine learning models.

**Week - 4**

This week includes –

Apply WordNet-based similarity measures on text.

Derive semantic topics from a large text collection using LDA.

List and describe techniques for named entity recognition and other information

extraction tasks.