

## **UC Merced**

### **Proceedings of the Annual Meeting of the Cognitive Science Society**

#### **Title**

Feature Visualization and Attribution Analysis of Confusion for Massive Open Online Course

#### **Permalink**

<https://escholarship.org/uc/item/66n3x2p1>

#### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

#### **Authors**

Zhao, Tongyu

Gao, Jiaying

Feng, Yu

et al.

#### **Publication Date**

2023

#### **Copyright Information**

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Feature Visualization and Attribution Analysis of Confusion for Massive Open Online Course

**Tongyu Zhao**

Universidade do Minho, Guimaraes, Braga, Portugal

**Jiaying Gao**

School of Artificial intelligence, changchun, Jilin province, China

**Yu Feng**

University of Minho, Guimarães, Portugal

**Jian Li**

College of Computer Science and Technology, Changchun, China

**Yatong Zu**

Jilin University, Changchun, China

**Sandro Pinto**

Universidade do Minho, Guimaraes, Braga, Portugal

**Adriano Tavares**

University of Minho, Braga, Portugal

**Hao Xu**

Jilin University, Changchun, China

## Abstract

Thousands of learners are taking massive Open Online courses (MOOCs) in search of self-improvement opportunities. However, the learning process often involves a struggle with confusion, which can adversely affect the course participation experience, leading to dropout halfway through. At the same time, it is impossible for the instructors to recognize and respond “confusion” posts accurately and timely, due to the large volume of registrants. This study aims to develop a robust and interpretable NLP model to identify “confusion” posts that require instructors’ immediate attention and to dig deeper into the origin causes of confusion posts according to the type of confusion. We find that posts identified as “confusion” are manifestations of different learner affects pertaining to their informational needs—primarily related to cognitive and social processes. Furthermore, we evaluated the model with a 13% improvement in F1 value compared to previous studies and successfully applied it across unseen domains.