europass

Alaettin Uçan

Date of birth: 04/02/1982 | **Nationality:** Turkish | **Gender:** Male | **Phone number:** (+90) 5456382878 (Mobile) | **Email address:** ucanalaettin@gmail.com

About me: I received my M.Sc.(2014) and Ph.D.(2020) degree from the Computer Engineering Department of Hacettepe University. The subject of my doctoral dissertation is the “Use of Optimization and Pretrained Language models in Emotion Analysis”. My main research topics are Information Retrieval, Natural Language Processing, Text Mining, Machine Learning, and Optimization. I am also a full-stack machine learning engineer.

WORK EXPERIENCE

01/03/2021 – CURRENT Osmaniye, Turkey
PROFESSOR OF COMPUTER ENGINEERING KORKUT ATA UNIVERSITY

I work as an Assistant Professor at Korkut Ata University, Department of Computer Engineering.

Business or Sector Education | **Department** Computer Engineering | **Address** Osmaniye, Turkey |

Website https://akbis.osmaniye.edu.tr/@aucan

31/01/2012 – 28/02/2021 Ankara, Turkey
RESEARCH ASSISTANT IN UNIVERSITY HACETTEPE UNIVERSITY

I worked as a research assistant at Hacettepe University Computer Engineering Department. In addition to teaching assistant duties in undergraduate and graduate courses, I took part in laboratory applications of courses such as programming, algorithm design, operating systems, software engineering, web services, database management systems, natural language processing, machine learning.

I worked as a researcher in many projects supported by TUBITAK (The Scientific and Technological Research Council of Turkey) at "Hacettepe University Multimedia Information Retrieval Laboratory". I took part in projects in Natural Language Processing, Text Mining, Machine Learning, Pre-trained language models, Transfer Learning, and Optimization research areas. I have written original research articles as a result of the projects we have developed both individually and as a team.

Business or Sector Professional, scientific and technical activities | **Department** Computer Engineering Dept. |

Address Ankara, Turkey | **Email** info@cs.hacettepe.edu.tr | **Website** http://cs.hacettepe.edu.tr/

31/05/2008 – 19/08/2010 Mersin, Turkey
SOFTWARE DEVELOPER KARMED SOFTWARE LTD.

I worked as a software developer in the Health Information Systems project developed within KarMed R&D. I have developed software modules for units such as clinic, surgery, warehouse, payment, and accounting in hospitals. The software modules that I developed using object-oriented languages in n-tier architecture have been used by thousands of users in hundreds of hospitals.

Department R&D | **Address** Mersin, Turkey | **Website** https://kardelensw.com/en/

EDUCATION AND TRAINING

31/01/2015 – 13/12/2020 Ankara, Turkey
PHD IN COMPUTER ENGINEERING Hacettepe University

Address Ankara, Turkey | **Website** http://cs.hacettepe.edu.tr/

31/01/2012 – 11/12/2014 Ankara, Turkey
MSC IN COMPUTER ENGINEERING Hacettepe University

Address Ankara, Turkey | **Website** http://cs.hacettepe.edu.tr/

11/09/2002 – 23/06/2007 Bishkek, Kyrgyzstan
BSC IN COMPUTER ENGINEERING KTU Manas

Address Bishkek, Kyrgyzstan | **Website** http://intl.manas.edu.kg/en

LANGUAGE SKILLS

Mother tongue(s): **TURKISH**

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
ENGLISH	C1	C2	B2	B2	C1
RUSSIAN	A2	A2	A2	A2	A1

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

DIGITAL SKILLS

Programming

Git | SQL | Javascript | Docker | C# | Delphi | Visual Studio | C | Java | C++ | JSON | HTML | Flask | Jupyter Notebook | MongoDB | Anaconda | Object-Oriented Programming | Github | MySQL | SQLite | Postman | PostgreSQL | NoSQL | Shell scripting (BASH) | Clean Code | Relational databases | Backend Developer | Regular Expressions | Visual Studio Code | Functional Programming

Machine Learning

Python | WEKA | Matplotlib | Tensorflow | Numpy | Deep Learning | Keras | Colab | Gensim | Django | Huggingface | Cuda | BERT | Feature Engineering | Transfer Learning | Pandas | SpaCy | LSTM | GRU | Scikit-Learn | CNN | NLTK | PyTorch | LaTeX | Natural Language Processing | Data Visualization | Data Preprocessing

Other

Microsoft Word | Microsoft Excel | Microsoft Powerpoint | Microsoft Teams

ADDITIONAL INFORMATION

PUBLICATIONS

Slime Mould Algorithm: A Comprehensive Survey of Its Variants and Applications – 2023

Meta-heuristic algorithms have a high position among academic researchers in various fields, such as science and engineering, in solving optimization problems. These algorithms can provide the most optimal solutions for optimization problems. This paper investigates a new meta-heuristic algorithm called Slime Mould algorithm (SMA) from different optimization aspects. The SMA algorithm was invented due to the fluctuating behavior of slime mold in nature. It has several new features with a unique mathematical model that uses adaptive weights to simulate the biological wave. It provides an optimal pathway for connecting food with high exploration and exploitation ability. As of 2020, many types of research based on SMA have been published in various scientific databases, including IEEE, Elsevier, Springer, Wiley, Tandfonline, MDPI, etc. In this paper, based on SMA, four areas of hybridization, progress

Farhad Soleimanian Gharehchopogh, Alaettin Ucan, Turgay Ibrikci, Bahman Arasteh, Gultekin Isik

A Study of Turkish Emotion Classification with Pretrained Language Models – 2022

Emotion classification is a research field that aims to detect the emotions in a text using machine learning methods. In traditional machine learning methods, feature engineering processes cause the loss of some meaningful information, and classification performance is negatively affected. Additionally, the success of modeling using deep learning approaches depends on the sample size. More samples are needed for Turkish due to the unique characteristics of the language. However, emotion classification datasets in Turkish are quite limited. In this study, the pretrained language model approach was used to create a stronger emotion classification model for Turkish. Well-known pretrained language models were fine-tuned for this purpose. The performances of these fine-tuned models for Turkish emotion classification were comprehensively compared with the performances of traditional machine learning and deep learning

methods in experimental studies. The proposed approach provides state-of-the-art performance for Turkish emotion classification.

Uçan A., Dörterler M., Sezer E. A.

An Emotion Analysis Scheme Based on Gray Wolf Optimization and Deep Learning – 2021

Automatic detection of emotions in textual data masses provides priceless opportunities for researchers and also it is inevitable for practitioners. The unfavorable factors involved in text data cause ambiguity and adversely affect the performances of emotion classifiers. Although deep learning approaches spark off significantly successive results, the obtained performances of the classifiers in the literature are commonly evaluated as the overall accuracy. This incomplete evaluation ignores inner class performance and overall accuracy can behave as a hopeful evaluator. In this study, we employed deep learning and meta-heuristic optimization methods together in order to resolve the ambiguity issue. Moreover, the decision mechanism of a conventional deep learning model is equipped with optimal emotion vectors obtained by optimization processes for each emotion class. Experimental results show that the proposed approach improves the inner class performance by maintaining the overall accuracy scores.

Uçan A., Dörterler M., Sezer E. A.

Emotion analysis in Turkish computational linguistics studies – 2020

Computational linguistics is an interdisciplinary field that aims to understand the verbal or written language, to express it mathematically and suggest methods, models and tools to achieve these goals. Emotion analysis, a research area of computational linguistics; is the process of finding which feelings taking place in what proportion in sound, image or text data. The developments such as the proliferation of the Internet, the increase of digital content, the increase of storage and computing power have both paved the way for automatic emotion analysis and made emotion analysis an important need. This study, which summarizes the subject of emotion analysis, aims to explain the history and importance of emotion analysis from a linguistic perspective, and briefly introduce the current application areas of emotion analysis.

UÇAN, A., SEZER, E. A.

A New Approach on Emotion Analogy by Using Word Embeddings – 2019

According to the Plutchik emotion classification, complex emotions consist of different combinations of eight basic emotions. The word embeddings in the literature are described as vector space where the word meanings are represented numerically. In this space, word analogies dealing with the similarities of vectors can be carried out. In this study, “emotion analogy” is proposed as a new method to create complex emotion vectors in case there is no learning data for complex emotions. In this respect, 12 complex feeling vectors were obtained by combining the word vectors of the basic emotions by the purposed method. The similarities between the obtained combinational vectors and the word vectors belonging to the complex emotions were investigated. As a result of the experiments performed on GloVe and Word2Vec word embeddings, it is found that the results of word analogy and emotion analogy are similar at 0.82 on average.

Uçan, A., & Sezer, E. A.

HUMIR at IEST-2018: Lexicon-Sensitive and Left-Right Context-Sensitive BiLSTM for Implicit Emotion Recognition – 2018

This paper describes the approaches used in HUMIR system for the WASSA-2018 shared task on the implicit emotion recognition. The objective of this task is to predict the emotion expressed by the target word that has been excluded from the given tweet. We suppose this task as a word sense disambiguation in which the target word is considered as a synthetic word that can express 6 emotions depending on the context. To predict the correct emotion, we propose a deep neural network model that uses two BiLSTM networks to represent the contexts in the left and right sides of the target word. The BiLSTM outputs achieved from the left and right contexts are considered as context-sensitive features. These features are used in a feed-forward neural network to predict the target word emotion. Besides this approach, we also combine the BiLSTM model with lexicon-based and emotion-based features. Finally, we employ all models in the final system using Bagging ensemble method. We achieved macro F-measure value of 68.8 on the official test set and ranked sixth out of 30 participants.

Naderalvojoud, B., Ucan, A., & Sezer, E. A.

SentiWordNet for New Language: Automatic Translation Approach – 2016

This paper proposes an automatic translation approach to create a sentiment lexicon for a new language from available English resources. In this approach, an automatic mapping is generated from a sense-level resource to a wordlevel by applying a triple unification process. This process produces a single polarity score for each term by incorporating all sense polarities. The major idea is to deal with the sense ambiguity during the lexicon transfer and provide a general sentiment lexicon for languages like Turkish which do not have a freely available machine-readable dictionary. On the other hand, the translation quality is critical in the lexicon transfer due to the ambiguity problem. Thus, this paper also proposes a multiple bilingual translation approach to find the most appropriate equivalents for the source language terms. In this approach, three parallel, series and hybrid algorithms are used to integrate the translation results. Finally, three lexicons are achieved for the target language with different sizes. The performance of three lexicons is evaluated in the lexicon-based sentiment classification task and compared with the results achieved by the supervised approach. According to experimental results, the proposed approach can produce reliable sentiment lexicons for the target language.

Ucan, A., Naderalvojoud, B., Sezer, E. A., Sever, H.

Imbalanced text categorization based on positive and negative term weighting approach – 2015

Although term weighting approach is typically used to improve the performance of text classification, this approach may not provide consistent results while imbalanced data distribution is available. This paper presents a probability based term weighting approach which addresses the different aspects of class imbalance problem in text classification. In this approach, we proposed two term evaluation functions called as PNF and PNF^2 which can produce more influential weights by relying on the imbalanced data sets. These functions can determine the significance of a term in association with a particular category. This is a crucial point because in one hand a frequent term is more important than a rare term in a particular category according to feature selection approach, and on the other hand a rare term is no less important than a frequent term based on idf assumption of traditional term weighting approach. Incorporation of these two approaches at the same time is the main idea that make them superior to other weighting methods. The achieved results from experiments which were carried out on two popular benchmarks Reuters-21578 and WebKB demonstrate that the probability based term weighting approach yields more consistent results than the other methods on the imbalanced data sets.

Naderalvojoud, B., Sezer, E. A., Ucan, A.

Assessment of feature selection metrics for sentiment analyses: Turkish movie reviews – 2014

Sentiment analysis systems pursuit the goal of detecting emotions in a given text with machine learning approaches. These texts might include three kinds of emotions such as positive, negative and neutral. Entertainment oriented texts, especially movie reviews, contain huge amount of possible emotional information. In this study, we aimed to represent each movie reviews by using small number of features. For this purpose, information gain, chi-square methods have been implemented to extract features for decreasing costs of calculations and increasing success rate. In experiments, employed corpus includes Turkish movie reviews, support vector machine and naïve bayes had been employed for classification and F1 score was used for performance evaluation. According to the experimental results, support vector machine achieved 83.9% performance value while classification of movie reviews in two (positive and negative) categories and also we obtained the 63.3% performance value while classification with support vector machine into three categories.

Akba, F., Uçan, A., Sezer, E. A., Sever, H.

PROJECTS

01/07/2021 – 01/07/2022

Automatic Question Generation Project I led a project for the automatic generation of knowledge-based multiple choice questions for a government agency.

I used keyword determination and summary methods with pre-trained language models to determine the information to be asked. Then I fine-tune pre-trained language models such as BERT and T5 to suit the question generation task. I fine-tune a separate model for paraphrasing the questions. I used word-embeddings to create multiple choice answer options. I was able to offer machine learning models as a service.

Within the scope of the successfully completed project, I orchestrated many technologies such as Python, Numpy, Pandas, Transformers, PyTorch, Gensim, WordNet, Word2Vec, BertTopic, BERT, T5, Django, PostgreSQL and many more.

30/06/2020 – CURRENT

YZYap: An educational robot project for beginners in programming and artificial intelligence. An educational robot project for beginners in programming and artificial intelligence. In addition to programming languages, Strach tool can also be used to program the robot. The robot can communicate with the user thanks to its microphone and camera sensors. In addition, the project provides a programmable and expandable artificial intelligence interface. The user can add artificial intelligence functions such as text classification, sentiment analysis, object tracking, and image recognition to the robot.

30/04/2017 – CURRENT

A project to monitor printing machines, personnel, paints and papers in a large printing house. A project to monitor printing machines, personnel, paints and papers in a large printing house. Thanks to the automation realized, production efficiency is increased by calculating input, output, personnel and machine statistics. On the other hand, the productivity of machinery and personnel is monitored and an alarm is given to the management in case of anomaly.

31/05/2017 – 13/12/2020

Emotion Classification with Pretrained Language Models The most successful results in the field were obtained by transferring prior knowledge in pre-trained language models, which is the state-of-the-art technique in natural language processing, to emotion analysis.

Link https://www.doi.org/10.1177/0165551520985507

31/10/2015 – 31/08/2018

Zigbee Based Wireless Smart Home Project I designed and developed a system for smart home equipment with Zigbee protocol developed by Gesislab Elektronik. I developed a web-based application that works with a system setup similar to Google Home and Amazon Alexa.
I have successfully worked in a complex environment where many technologies such as C, C++, Python, C#.Net, PostgreSQL, NoSQL, Javascript are used together.

14/03/2016 – 14/06/2017
Turkish Rule Based Sentiment Analysis System Although the supervised approach performs well on the sentiment classification task, the availability of sentiment annotated data is known as a limitation for this approach. On the other hand, the term-based features like bag-of-words or n-grams cannot make more progress on the performance of this approach in cases that the sentiments of several texts are presented by more ambiguous words or phrases. This is important because natural language is ambiguous. In this condition, the sentiment lexicons play important role in sentiment analysis systems. This is considerable for supervised approach since these lexicons can be used in extracting more effective features along with term-based ones. However, despite the successful performance of using these lexicons in English sentiment analysis systems, they cannot be employed in a new language due to the lack of such lexical resources. This project proposes an automatic translation approach to create a sentiment lexicon for a new language from available English resources. In this approach, an automatic mapping is generated from a sense-level resource to a word-level by applying a triple unification process. This process produces a single polarity score for each term by incorporating all sense polarities. The major idea is to deal with the sense ambiguity during the lexicon transfer and provide a general sentiment lexicon for languages like Turkish which do not have a freely available machine-readable dictionary. On the other hand, the translation quality is critical in the lexicon transfer due to the ambiguity problem. Thus, this project also proposes a multiple bilingual translation approach to find the most appropriate equivalents for the source language terms. In this approach, three parallel, series and hybrid algorithms are used to integrate the translation results. Finally, three lexicons are achieved for the target language with different sizes. The generated lexicons are used in a rule-based sentiment classification process and compared with the supervised approach.

01/01/2017 – 29/03/2017
Ford - Fiberlast Laser Control Automation I designed and built the management and communication interface to the laser marking device produced by a local company for use in the Ford Sakarya factory.

I decoded the Ford production communication protocol and extracted the necessary information and saved it to the database. I forwarded the information waiting in line to be marked in the database to the laser device. In addition to the marking device, I successfully managed the auxiliary equipment such as the label coil drive device, the label cutting device and the smoke absorber.

My application, which I use C++ language as well as MySQL and C#.Net technologies, has been working without any problems for years.

01/01/2017 – 31/01/2017
Hacettepe n-Gram Extractor I created an application that can perform text mining and natural language operations to be used in research within the Department of Linguistics, Turcology and Dictionary Research Center.
I developed the application in Python using libraries such as Numpy, Pandas, NLTK, Zemberek.

30/09/2014 – 30/09/2015
Written media review and analysis system I built a print media tracking system for a government agency. Thanks to the system, predetermined labels are highlighted after scanning and OCR processing of daily newspapers and magazines. Operators can create daily news bulletins by checking these highlighted sections and cutting the relevant clippings.
I actively used C#.Net, Abbyy Fine Reader API, MSSQL, Javascript, AJAX technologies.

01/01/2013 – 30/09/2013
A third party filter to Metastock trading platform I developed a 3rd party Stock Star filter that works on the Metastock platform. For this filter; I developed licensing, payment system, virtual POS, installation package, training and help modules.
I developed a web application and a DLL using C#.Net and C++ languages.

31/05/2008 – 19/08/2010
Health Information Management Systems I took an active role in the development of the Hospital Information System software, which is used by thousands of users in more than two hundred hospitals across the country.
Service architectures, database design, user interfaces and CRUD operations were among my responsibilities.
I have developed more than 10 modules with Delphi, Oracle, C#.Net and C++.
I have worked with protocols such as HL7 and ASTM.
I have successfully performed data consistency and query optimization.