Ali Najafi
najafi-ali.com

GitHub: @alinajafi1998
Email: ali.najafi@sabanciuniv.edu

RESEARCH INTERESTS

My research spans the interdisciplinary fields of machine learning and natural language processing, with a particular focus on addressing societal challenges through computational methods. I aim to develop tools and apply techniques to uncover complex patterns that can be leveraged to tackle pressing social issues.

EDUCATION

Sabanci University, Istanbul, Turkey   *Feb. 2022 - May. 2024*
M.Sc., Computer Science and Engineering
Thesis: Developing Turkish Language Models on Social Media
Advisor: Asst. Prof. Onur Varol
CGPA: 3.71/4.0

University Of Tabriz, Tabriz, Iran   *Sept. 2017 - Sept. 2021*
B.Sc., Computer Science and Engineering
CGPA: 16.58/20

PUBLICATIONS
GOOGLE SCHOLAR
CITATIONS: 18
H-INDEX: 3

Journal Articles in preparation

J.5 **Najafi, A.** and { Alp Onder Yener, Letizia Iannucci, Mikko Kivelä, Onur Varol } **"Social Media Data Adaptors: Standardization and Anonymization of Social Network Datasets"**

J.4 **Najafi, A.**, and Onur Varol. **"Analyzing Hateful Content During the Nights of the 2023 Turkish Election"**

Journal Articles

J.3 **Najafi, A.**, & Varol, O. (2024). **"Turkishbertweet: Fast and reliable large language model for social media analysis."** Expert Systems with Applications, 255, 124737. ↗

J.2 **Najafi, A.**, Mugurtay, N., Zouzou, Y., Demirci, E., Demirkiran, S., Karadeniz, H. A., & Varol, O. (2024). **"First public dataset to study 2023 Turkish general election."** Scientific Reports, 14(1), 8794. ↗

J.1 **Najafi, A.**, Gholipour-Shilabin, A., Dehkharghani, R., Mohammadpur-Fard, A., & Asgari-Chenaghlu, M. (2023). **"ComStreamClust: a communicative multi-agent approach to text clustering in streaming data."** Annals of data science, 10(6), 1583-1605. ↗

Refereed Conference Proceedings

C.2 Yener, A. Ö., **Najafi, A.**, & Varol, O. (2024, May). **"City-level User Location Detection from Social Media and Analysis of Political Participation."** In 2024 32nd Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE. ↗

C.1 **Najafi, A**, & Varol, O. (2024, March). **"VRLLab at HSD-2Lang 2024: Turkish hate speech detection online with TurkishBERTweet."** In Proceedings of the 7th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2024) (pp. 185-189). ↗

CONDUCTED RESEARCH AND ACADEMIC EXPERIENCE

VRL LAB, Istanbul, Turkey   *Sept 2022 - Present*

- **TurkishBERTweet**
  * Achieved almost 10% significant improvement in the F1-Score of Text Classification Tasks
  * Developed classification pipelines for Sentiment Analysis and Hate Speech Detection Tasks
  * Trained TurkishBERTweet From Scratch

* Used Google TPUs for Training the Language Model
* Trained BPE tokenizer from scratch on 894M Turkish Twitter Data
  - **Turkish Election Dataset**
    * Analyzed the Retweet Speed of specific Turkish Politicians for different time spans of election day
    * Conducted network analysis of politicians on Twitter using Networkx and Gephi
    * Analyzed the followers of politicians to identify social bots using Botometer
  - **CON-NET** under CHIST-ERA call funded by EU and TUBITAK
    * Developed a Tool for Standardizing and Anonymizing Social Network Datasets

---

INVITED
LECTURER AND
TALKS

1. Presented the paper "VRLLab at HSD-2Lang 2024: Turkish Hate Speech Detection Online with TurkishBERTweet" at EACL"2024 📄 , Malta, March 2024.

2. Presented the paper "First Public Dataset to Study the 2023 Turkish General Election" at the Turkey Computational Social Sciences Conference (TCSS) 2023 📄 , Koc University, Turkey.

3. Invited lecturer for the talk "Applications of Large Language Models for Social Scientists" in the Political Science course at Sabanci University, Turkey, under the supervision of Dr. Fatih Serkant Adiguzel, Fall 2023.

4. Invited lecturer for the talk "Natural Language Processing for Social Scientists" at the Summer Institutes in Computational Social Science (SICSS) 📄 , Ozyegin University, Turkey, Summer 2023.

TEACHING
EXPERIENCE

Sabancı University - Faculty of Engineering and Natural Sciences
CS210: Introduction to Data Science *Spring 2023, Fall 2023*
CS412: Machine Learning *Fall 2024, Spring 2024, Fall 2022*
CS201: Introduction to Computing *Spring 2022*

Held weekly office hours and recitation sessions to address course-related questions.
Graded exams and assignments

COMPUTER AND
LANGUAGE
SKILLS

Skills:
    Programming Languages: Python, C++, Java
    Frameworks: Pytorch, Tensorflow, Keras, Flax, Django
    Librares: Matplotlib, Seaborn, Pandas, Numpy, Gephi, NetworkX, OpenCV, Jax, Docker
    DataBase: SQL, NoSQL (Redis)
    Familiar with GCP, AWS, LaTeX, HTML, CSS, JS

Languages: English (Fluent), Persian (Native), Azerbaijani (Native), Turkish (Fluent)

HONORS AND
AWARDS

Hate Speech Detection in Turkish - EACL"2024 *Jan. 2024*
Earned **3rd** Rank among 11 participated teams

TUBITAK Scholarship *Aug. 2023 - Present*
The Scientific and Technological Research Council of Turkey, Turkey
Awarded a full tuition waiver and monthly stipend as a researcher on a TUBITAK-funded project.

Sabancı University Scholarship *Feb. 2022 - Aug 2023*
Sabancı University, Istanbul, Turkey
Awarded a 100% tuition scholarship and a monthly stipend as a Teaching Assistant