

Mert Yuksekgonul

Ph.D. Student in Computer Science, admitted Autumn 2021

Publications

PUBLICATIONS

- **Optimizing generative AI by backpropagating language model feedback.** *Nature*
Yuksekgonul, M., Bianchi, F., Boen, J., Liu, S., Lu, P., Huang, Z., Guestrin, C., Zou, J.
2025; 639 (8055): 609-616
- **How Well Can LLMs Negotiate? NEGOTIATIONARENA Platform and Analysis**
Bianchi, F., Chia, P., Yuksekgonul, M., Tagliabue, J., Jurafsky, D., Zou, J.
edited by Salakhutdinov, R., Kolter, Z., Heller, K., Weller, A., Oliver, N., Scarlett, J., Berkenkamp, F.
JMLR-JOURNAL MACHINE LEARNING RESEARCH.2024
- **A visual-language foundation model for pathology image analysis using medical Twitter.** *Nature medicine*
Huang, Z., Bianchi, F., Yuksekgonul, M., Montine, T. J., Zou, J.
2023
- **GPT detectors are biased against non-native English writers.** *Patterns (New York, N.Y.)*
Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., Zou, J.
2023; 4 (7): 100779
- **Meaningfully Debugging Model Mistakes using Conceptual Counterfactual Explanations**
Abid, A., Yuksekgonul, M., Zou, J.
edited by Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., Sabato, S.
JMLR-JOURNAL MACHINE LEARNING RESEARCH.2022: 66-88