

Nalin KUMAR

Master's Student: Computer Science

[in linkedin.com/in/nalin-kumar-7b92aa174](https://www.linkedin.com/in/nalin-kumar-7b92aa174) github.com/knalin55
☎ +420 775 171 703 @ knalin55@gmail.com
📍 Prague, Czech Republic
[🔗 Personal Webpage](#) [🔗 Google Scholar](#)

EDUCATION

2024 - ongoing	Institute of Formal and Applied Linguistics, MFF CHARLES UNIVERSITY PhD in Computational Linguistics Thesis Advisor: Mgr. Ondřej Dušek, Ph.D. Thesis title: Efficient Methods for Natural Language Generation Systems	Prague, Czech Republic
2021 - 2024	Faculty of Mathematics and Physics CHARLES UNIVERSITY Master's in Computer Science - Language Technologies and Computational Linguistics Thesis Advisor: Mgr. Ondřej Dušek, Ph.D. Thesis title: Neural Models for Multilingual Natural Language Generation and Understanding	Prague, Czech Republic
2016 - 2021	National Institute of Science Education and Research (NISER) HBNI Integrated Master's (7.41/10) Thesis Advisor: Dr. Subhankar Mishra Thesis title: Neural Machine Translation for Low Resource Languages Major Mathematics Minor Computer Science	Bhubaneswar, India

SCHOLARSHIPS

2021-23	Computational Linguistics Scholarship (CLS 2021) CUNI, Czech Republic
2016-21	INSPIRE Scholarship DST, India

PUBLICATIONS

2024	LEETs-Dial: Linguistic Entrainment in End-to-End Task-oriented Dialogue systems [Link] Nalin Kumar, Ondřej Dušek Findings of the Association for Computational Linguistics: NAACL 2024
2023	Better Translation + Split and Generate for Multilingual RDF-to-Text [Link] Nalin Kumar, Saad Obaid ul Islam, Ondřej Dušek Proceedings of the Workshop on Multimodal, Multilingual Natural Language Generation and Multilingual WebNLG Challenge (MM-NLG 2023)
2022	Genre Transfer in NMT: Creating Synthetic Spoken Parallel Sentences using Written Parallel Data [Link] Nalin Kumar, Ondřej Bojar Proceedings of the 19th International Conference on Natural Language Processing (ICON)
2021	QUARC: Quaternion Multi-Modal Fusion Architecture For Hate Speech Classification [Link] Deepak Kumar*, Nalin Kumar*, Subhankar Mishra 2021 IEEE International Conference on Big Data and Smart Computing (BigComp) *equal contribution
2021	NLP@NISER: Classification of COVID19 tweets containing symptoms [Link] Deepak Kumar*, Nalin Kumar*, Subhankar Mishra Proceedings of the Sixth Social Media Mining for Health (SMM4H) Workshop and Shared Task, NAACL 2021 *equal contribution

Dict-NMT: Bilingual Dictionary based NMT for Extremely Low Resource Languages [arXiv]
Nalin Kumar, Deepak Kumar, Subhankar Mishra

WORK EXPERIENCE

Nov 2023 Sept 2024	Research Assistant CZECH TECHNICAL UNIVERSITY Abstractive text summarization of scientific texts.	Prague, Czechia
Nov 2022 Feb 2023	Research Assistant CHARLES UNIVERSITY Humane AI Microproject CUNI/BUT – Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems.	Prague, Czechia
Aug 2020 Oct 2020	Data Science Intern TATA STEEL Worked under Mr. Sanjoy Biswas, Head of Advanced Analytics, Kalinganagar and Mr. Saibal Nandi, Head Analytics Marketing and Sales. <ul style="list-style-type: none">> Worked on detecting and counting steel reinforcement bars using deep learning.> Worked on Dealer names matching (Entity Matching) using Machine Learning.> Worked on Document Verification for FAMD Exports, TATA Steels, using text analytics and its UI development (using Streamlit). It got deployed on GCP so that the app is accessible to the users remotely during COVID-19 times. The project is identified under the Shikhar Initiative and is estimated to save approx. INR 1 Cr.	India
Dec 2020	Winter Research Intern JNU 2 weeks winter internship under Dr. Narinder Singh Sahni on Gaussian process.	New Delhi, India
May 2019 Jul 2019	Summer Research Intern JNU Summer research internship under Dr. Narinder Singh Sahni on Regression Analysis. <ul style="list-style-type: none">> Parametric Regression models> Explored different simpler linear and non-linear parametric models arising from the five-point non-linear model used by the vendors to fit the concentration-absorbance curve of HbA1c(used in the measurement of average sugar levels over a period of 3 months).	New Delhi, India
Jun 2018	Summer Research Intern IIT HYDERABAD Summer internship under Dr. Narasimha on Introduction to Cryptography.	India
Jul 2017 May 2017	Summer Research Intern HARISH CHANDRA RESEARCH INSTITUTE (HRI) Summer internship under Prof. Kalyan Chakraborty on Group Theory.	Allahabad, India

UNIVERSITY PROJECTS

2021 present	Course Projects CHARLES UNIVERSITY Projects taken and completed during the educational stay at Charles University in reverse chronological order. <ul style="list-style-type: none">> (2022-23) Implemented an existing Information Retrieval framework, PyTerrier, for topic retrieval task.> (2021-22) Implemented a task-based dialogue system with PyTorch-lightning.> (2021-22) Implemented an HTK toolkit for ASR of manually created Hindi dataset.> (2021-22) Worked on a survey report on "Bilingualism: A study on the contact of two linguistic systems" for a Linguistics course.	Prague, Czechia
-----------------	--	-----------------

2016	Course Projects NISER	Bhubaneswar, India
2021	Projects taken and completed during the educational stay at NISER in reverse chronological order. <ul style="list-style-type: none"> > (2020-21) Reviewed existing works and implemented many-to-one NMT systems for low-resourced languages. > (2020-21) Neural Machine Translation of Vedic Sanskrit to English > (2020-21) Hateful Meme Challenge by DrivenData > (2019-20) Hate Speech Classification on Twitter Dataset > (2019-20) Emotion detection of Tom & Jerry. > (2019-20) COVID-19 Face-mask Detection. 	

WORKSHOPS, PROGRAMMES & VOLUNTEERING EXPERIENCES

2023	Advanced Language Processing School (Winter School) REMOTE Presented a poster on "Exploring joint approaches to RDF-based semantic parsing".	France
2020	SciPy India 2020 (Workshop) REMOTE Conducted a 2 hour short workshop on "Object counting in an image via Object Detection and Marking"	Bhubaneswar, India

SKILLS

Programming Languages	Python R Bash
Deep Learning Frameworks	PyTorch Tensorflow Keras HuggingFace APIs and Libraries
Misc	Slurm & SGE Streamlit MATLAB MS VS Code

REFERENCE(S)

Mgr. Ondřej Dušek, Ph.D.
Assistant Professor
 Institute of Formal and Applied Linguistics, MFF, Charles University
 odusek@ufal.mff.cuni.cz

doc. RNDr. Ondřej Bojar, Ph.D.
Associate Professor
 Institute of Formal and Applied Linguistics, MFF, Charles University
 bojar@ufal.mff.cuni.cz

Dr. Subhankar Mishra
Reader-F
 School of Computer Sciences, NISER, India
 smishra@niser.ac.in