

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

2021 - 2023	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

2022	Genre Transfer in NMT: Creating Synthetic Spoken Parallel Sentences using Written Parallel Data Nalin Kumar, Ondřej Bojar Accepted at 19th International Conference on Natural Language Processing (ICON 2022)
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 Accepted as a shared task system description at 6 th SMM4H, NAACL 2021 *equal contribution
In review	Dict-NMT: Bilingual Dictionary based NMT for Extremely Low Resource Languages [arXiv] Nalin Kumar, Deepak Kumar, Subhankar Mishra
In work	Shortening MT outputs using paraphrasing systems Nalin Kumar, Dávid Javorský, Ondřej Bojar

WORK EXPERIENCE

Nov 2022 Feb 2023	Research Assistant CHARLES UNIVERSITY Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems.	Prague, Czechia
----------------------	--	-----------------

Oct 2020 Aug 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 detection and counting of 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) which is further deployed on GCP so that the app is accessible to the users remotely in COVID times. The project is identified under 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 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 2023	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. ➤ (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 2021	Course Projects NISER 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. 	Bhubaneswar, India

WORKSHOPS & PROGRAMMES

2020

SciPy India 2020| REMOTE

Bhubaneswar, India

Conducted a 2 hour workshop on "Object counting in an image via Object Detection and Marking"

SKILLS

Programming Languages	Python R \LaTeX 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.

Associate Professor

Institute of Formal and Applied Linguistics, MFF, Charles University

odusek@ufal.mff.cuni.cz

doc. RNDr. Ondřej Bojar, Ph.D.

Assistant 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