SAHIL CHOPRA

+33 751851792 | sahilc.mac@gmail.com | schopra6.github.io

linkedin.com/in/schopra410 | Nancy,France

EDUCATION

Universitat des Saarlandes	Saarbrucken ,DE
Masters of Science in Language Science and Technology	*Oct 2022 -– Sept 2023*

- Part of Erasmus Mundus Language, Communication and Science program

Université de Lorraine	Nancy ,FR
Masters of Science in Natural Language Processing	*Sept 2021 -– Aug 2022*

- Part of Erasmus Mundus Language, Communication and Science program

Guru Gobind Singh Indraprastha University	Delhi, IN
Bachelor of Technology in Information Technology	*July 2015 -– May 2019*

- Cumulative GPA 8.9/10
- Relevant Coursework: Linear Algebra, Calculus, Statistics, Big Data Analytics, Probability

RESEARCH EXPERIENCE

Synalp Lab, Loria	Oct,2021 —
Research student under supervision of Dr. Claire Gardent	*Nancy, Fr*

- Working on cross-lingual and multilingual approach in natural language generation (NLG).The aim is to perform summarisation, simplification and data-to-text generation using NLG.

MIDAS Lab, IIIT Delhi	June,2018 — Sept,2018
Research Intern under supervision of Dr. Rajiv Ratn	*Delhi, IN*

- Developed a multi-modal technique to predict popularity of a social media post under ACMMM challenge
- Worked on a model to estimate the locations where multimedia items were captured using an algorithm that inspects the content and metadata.
- Developed an automated feature-based SVM classification technique that can help selective tweets reach relevant medical authorities.

PROFESSION EXPERIENCE

United Health Group	July,2019 — Sept,2021
Software Engineer, Data Solution	*Gurugram, IN*

- Formulated a semantic role labelling network model for parsing insurance claims and process it for appeal team.
- Developed high-performant micro-services built on top of Node.js and Apollo GraphQL using Javascript to provide access control on streaming data .
- Developed features to generate extracts from massive data of claims submitted using Apache Spark , Hive and Scala which saved $40k annually.

PUBLICATION

- **S. Chopra**, A Gupta and A Kaushik. Knowledge Directed Multi-task Framework for Natural Language Inference in Clinical Domain *Proceedings of 18th ACL Workshop on BioNLP, ACL 2019*, Florence, Italy, pages 488–492.
- P. Mathur, M. Ayyar, **S. Chopra** , S. Shahid, L. Mehnaz and R. Shah . Identification of Emergency Blood Donation Request on Twitter. *Proceedings of the 2018 EMNLP Workshop SMM4H: The 3rd Social Media Mining for Health Applications Workshop Shared Task*, Brussels, Belgium, pages 27–31.

TECHNICAL SKILLS

Languages: Java, Python,Scala C/C++, SQL, Javascript ES6
Developer Tools: Git, Docker, Kubernetes, Azure, Jupyter, Intellij
Libraries: TensorFlow, Pytorch, pandas, NumPy, Matplotlib, OpenCV, GraphQL
Frameworks: Apache Spark, Apache Kafka, Apache Avro, Springboot, Django, Express

HONORS AND AWARDS

* Selected for prestigious Erasmus+ Scholarship to pursue Masters in Europe.
* Received multiple bravo awards in UHG for performance and innovation in projects.
* Select for the Inter University final year project competition among 120 students.
* Received AICTE grant to present paper in EMNLP conference.
* Runner up in Research Project presentation Techvilla,Avensis.

PROJECTS

Speech to Text Diarisation Tool   Oct,2021 —

The project is an application which takes two users voice, identify them and convert their speech into text. At the beginning, the application only treats the English language, but other languages may be added in further versions. It may also be able to recognize more than two speakers.

Wide CNN-MI Paraphrase Identification   April,2018 — Sept,2018

This involves a Convolution Neural Network based architecture for Paraphrase Identification comparing two sentences on multiple levels of granularity and learns model interaction features at each level.

Language Identification of Codemixed Text   April,2017 — June 2017

In this project, model first outputs a language distribution for every token independently with efficient feed-forward classifiers. Then, a decoder chooses labels using both the token predictions and global constraints over the entire sentence. This decoder produces high-quality predictions on monolingual texts as well as codemixed inputs.

POSITION OF RESPONSIBILITY

- Mentor for IEEE summer project , MSIT
- Event Head(Algojam) at Avensis Annual fest MSIT
- Department Head of UI/UX design at Ecell MSIT
- Event Coordinator at Avensis Annual fest MSIT