Sonal Sannigrahi
76 Boulevard des Maréchaux, Palaiseau- 91120
+33 0662574908 | sonal.sannigrahi@polytechnique.edu |**Github**: sonalsannigrahi

EDUCATION

École Polytechnique, Palaiseau, France   2018-2021
- BSc. in Mathematics and Computer Science | **GPA: 3.93/4.0**
- Admitted with **Honours** | Honours Grant of 10500 Euros | **Google WTM Scholar EMEA 2020**
- <u>Relevant Coursework:</u> Design and Analysis of Algorithms, Machine Learning, Computer Architecture, Intro to Networks, OOP in C++, Functional Programming, Probability and Statistics, Algebra, Analysis, Discrete Mathematics, Formal Languages, Deep Learning in Computer Vision (Graduate Course)

TECHNICAL EXPERIENCE and OUTREACH

Linagora, Paris, France- **Research and Development Intern**   Summer 2020
- Built an on-device next word predictor in French and English similar to mobile keyboards, for emails
- Testing and developing new architectures: Bi-LSTM, Transformer based, CNN-LSTM, GRU
- Comparing different word2vec representations to attain better accuracy on word level language models
- Used knowledge distillation to compress models, experimented with Flask for web deployment

École Polytechnique, Palaiseau, France- **Research Assistant**   Spring 2020
- Worked under Prof. Thomas Simonson to develop a Monte Carlo method to determine classes in a sequence alignment based on a probability distribution and train a MLP network to learn these positions

École Polytechnique, Palaiseau, France- **Research Assistant**   Summer 2019
- Worked under Prof. Pierre Boyer to write code in **Python** to calculate Twitter Penetration Rates for a given region and time period. Used **R** to create several different visualisation techniques at city, county, state levels of accuracy. Trained Naïve Bayes Classifier model for sentiment analysis on bilingual tweets; **presented at PyCON FR 2019 as** "APIs and Language Processing with Python for Twitter"
- Used Twitter REST API and Get OAuth methods to extract and geolocate twitter JSON data.

Girls in Tech and Code Club, Paris, France - *Volunteer Instructor*   Spring/Summer 2019
- Gave introductory Computer Science and Mathematics courses with a frequency of 2 hrs/week to a group of 30 + girls from middle and high school in Paris region.
- Created modules to teach different projects in **Scratch, Java Script + HTML/CSS** in terms of small interactive games and basic web development.

RaspberryPi Foundation, Remote - *Translation Volunteer*   Jul 2019 - Present
- Translated more than 10 projects by the RaspberryPi foundation in **JavaScript, HTML/CSS, Python, Scratch 3** and using the **RaspberryPi3** from English to Hindi. '
- Lead a team of 6 people to organize an online Translation Hackathon to complete over 15 projects over the span of 24 hours, reached nearly 250 code clubs in India.

SELECT PROJECTS

Hindi OCR System [Class Project for Deep Learning in Computer Vision, Grade: A, **Pytorch and OpenCV**]
- Built a novel OCR system using a deep convolutional neural network achieving 98.25% accuracy on character level recognition and 90% accuracy on paragraph level. Designed new algorithms for text preprocessing specifically for Hindi and other Devanagari based scripts
- **Accepted** into the 9th International Conference for Natural Language Processing, Zurich (Declined)

Neural Recommender System [Class Project for Machine Learning, Grade: A+, **Tensorflow**]
- Implemented a redesigned recommender system using Neural Matrix Factorisation (NeuMF) and compared against general MF achieving better results on Hit Ratio and NDCG metrics

Tic Tac Toe [Class Project for Networks, Grade: A+, **Coded in C**]
- Designed and Implemented client server UDP sockets to play with 4 functional clients + 2 player server

Archers [Class Project for OOP in C++, **Coded in C++**, view code on Github]

SKILLS AND ADDITIONAL INFORMATION

Technical: Python, C++, familiar with: HTML/CSS, JavaScript, R, Shell, OCaml
Languages: English, Hindi, French
Google Get Ahead 2020: Invite only development programme in EMEA, coding weekly assignments in Python
Reviewer at NLP Beyond Text Workshop EMNLP 2020: Part of PC to review papers on multimodal processing