Onkar Pandit

INRIA Lille - Nord Europe
40 Avenue du Halley, 59650
Villeneuve-d'Ascq, Lille, France

Phone : +33-768825850
email : onkar.pandit@inria.fr

Education

- *Doctor of Philosophy* - Computer Science   Dec. 2017 - present
  Title - Graph based semi-supervised linguistic structure prediction
  INRIA, Lille, France

- *Master of Technology* - Electrical Engineering (7.0/10.0)   Jul. 2012
  Specialization - Signal processing and communication networks
  Indian Institute of Technology, Kanpur

- *Bachelors of Technology* - Electronics and Telecommunications Engineering (8.37/10.0)   May 2010
  Shri Gurugobind Singhji Institute of Engineering and Technology, Nanded, Maharashtra

- *Higher Secondary Certificate* (87.33/100.0)   Jun. 2006
  Maharashtra State Board

- *Secondary School Certificate* (87.06/100.0)   Jun. 2004
  Maharashtra State Board

Experience

- Indian Statistical Institute, Kolkata — Kolkata, India
  - *Project Scientist*   Jul. 2016 – Nov. 2017

- Oracle India Pvt. Ltd. — Bangalore, India
  - *Senior Member Technical Staff*   Mar. 2016 – Jun. 2016

- Oracle India Pvt. Ltd. — Bangalore, India
  - *Member Technical Staff*   Jul. 2012 – Mar. 2016

- Indian Institute Of Technology, Kanpur — Kanpur, India
  - *Project Associate*   Jun. 2011 – Jul. 2012

- Indian Institute Of Technology, Kanpur — Kanpur, India
  - *Teaching Assistant*   Jul. 2010 – May 2011

Publications

- Chaturvedi A., **Pandit O.**, Garain U., *CNN for Text-Based Multiple Choice Question Answering*, Association for Computational Linguistics (ACL-2018).
- Chakrabarty A., **Pandit O.**, Garain U., *Context Sensitive Lemmatization Using Two Successive Bidirectional Gated Recurrent Networks*, Association for Computational Linguistics (ACL-2017).
- Garain U., **Pandit O.**, Kise K., Augereau O. *Identification of Reader Specific Difficult Words by Analyzing Eye Gaze and Document Content*, International Conference on Document Analysis and Recognition (ICDAR-2017).

Awards and Achievements

- Oracle webcenter best employee'16 *Numero Uno* award.
- Part of Oracle webcenter best performing team'16 award.
- Ministry of Human Resource Development, India, postgraduate study fellowship in 2010.
- Secured 310 rank out of 1,00,000+ aspirants in Graduate Aptitude Test in Engineering 2010.

Projects and Research

BioMedical Natural Language Processing   Jul. 2016 – Nov. 2017

- Worked on Biomedical Named Entity Recognition task.
- Designed a novel neural architecture and developed Python code to build it.
- Achieved new state-of-the-art result with 93% accuracy.
- Currently, working on Biomedical event extraction(BioEE) problem. Developing neural network based system for BioEE.

Code-mixed Social Media Natural Language Processing   Oct. 2016 – Nov. 2017
- Working on Code-mixed Social Media texts NER and POS tagging problem.
- Processed WhatsApp, Twitter and Facebook data for Hindi-English, Bangla-English, Tamil-English languages.
- Designed and developed a neural architecture for it. Working on it to achieve good results.

Reading Behavior Analysis   Jan. 2017 – Apr. 2017
- Involves prediction of difficult words in a document from a readers' eye gaze data.
- Jointly proposed SVM based model and used both eye gaze data information and linguistic aspects of document to set new state-of-the art performance at 62%.
- Developed a system based on it in Python.

Lemmatizer for languages   Aug. 2016 – Mar. 2017
- Task of finding a root word from the given word(*e.g. "went" to "go"*)
- Developed neural network based architecture for it and tested it against various languages such as Bengali, Catalan, Dutch, Hindi, Hungarian, Italian, Latin, Romanian and Spanish.
- Achieved new state-of-the-art results with proposed architecture with all languages except Bengali.

Oracle Big Data Preparation Cloud Service(BDP)   Sep. 2015 – Jun. 2016
- NLP based product, which processes ingestion, preparation, enrichment, and governance of data automatically.
- Contributed to the product development process since its conception.
- Worked as a part of infrastructure provisioning team.
- Developed various tools to enhance the productivity of team.

Oracle Webcenter Portal   Jul. 2012 - Aug. 2015
- Includes WebCenter Portal services like instant messaging, document and content management, discussion forums, wikis, blogs, tagging, activity tracking, personalization, polls, and others.
- Worked on the software development of Portal Security, Notifications, Notes, Relations, etc services.
- Improved test automation framework to catch defects effectively.

Detection of signal in cognitive radio with cyclostationarity (*M.Tech Thesis)*   May 2011 - Jun. 2012
- Cognitive Radio works on the principle of opportunistic communication.
- Detecting the vacant frequency band in cognitive radio is crucial step.
- Proposed a method for detection of signal using signal's cyclostaionarity property.

Relevant Coursework

- Machine Learning
- Deep Learning
- Probability and Statistics
- Linear Algebra
- Data Structures
- Algorithms

Skills
- *Programming Languages:* Python, Java, R and C
- *Software:* Sklearn, Keras, LaTeX
- *Operating Systems:* Linux/Unix system, Windows