7a.029.TUT - Very Fast Nearest Neighbor Retrieval in High-Dimensional Domains

Project - Team

Team Member	Role	Email	Phone Number	Academic Sites/Industry Members
Petri Myllymäki	Co-PI	petri.myllymaki@helsinki.fi	+358 40 5531162	University of Helsinki
Teemu Roos	PI	teemu.roos@cs.helsinki.fi	Not available	University of Helsinki
Jukka Kohonen	Researcher	jukka.kohonen@cs.helsinki.fi	Not available	University of Helsinki
Ville Hyvonen	Student	ville.hyvonen@cs.helsinki.fi	Not available	University of Helsinki
Janne Leppa-Aho	Student	Not Available	Not Available	University of Helsinki
Gur Ersalan	Student	Not Available	Not Available	University of Helsinki
Kimmo Valtonen	Project Mentor	kimmo.valtonen@m-brain.com	Not available	**Funded By: M-Brain**

Project - Summary

This project investigates variants of approximate nearest neighbor (ANN) search algorithms based on random projection trees.

The objectives are to develop and evaluate: i) scalable algorithms for ultrahigh-dimensional data by exploiting sparsity; ii) incremental algorithms for dynamic (streaming) data; iii) robust (fault-tolerant) parallelization techniques on GPU, cluster and cloud platforms.

Applications involve text, audio, image and other signal data.

Project - Novelty of Approach

- We use multiple random projection trees, which allows for a trade-off between the computational cost and the probability of retrieving the closest neighbor or neighbors

Project - Details of Progress/Achievement

We have applied MRPT (multiple random projection trees) method to a large-scale text classification task. MRPT made nearest-neighbor based classification possible as the exact search was computationally infeasible. However, nearest-neighbor classifier itself was outperformed by other methods in this setting.

We have made progress in research regarding the fault-tolerant parallelization techniques and in the automatic tuning of the parameters of the method.

Project - Deliverables

	Deliverable
1.	An open-source package for parallelized ANN.
2.	Concrete applications of ANN in large-scale document classification, clustering, etc.
3.	Theoretical framework for the study of parallel ANN and other machine learning procedures.
4.	Optional deliverable: Incremental algorithms for dynamic (streaming) data.

Attention Project PIs

PIs are responsible for keeping up the content of their project page and have the ability to EDIT the page.

- To **EDIT**, click the edit "pencil icon" in the top right-hand corner of this page
- To **PUBLISH** your changes, click the blue "Publish" button in the lower right-hand corner of this page
- If you need help or have questions, please contact Site Admin: Sally.Johnson@louisiana.edu

Table of Contents

Spaces

- All Spaces

CVDI 2017 IAB Fall Meeting

CVDI 2018 IAB Fall Meeting

CVDI 2018 IAB Spring Meeting

CVDI 2019 IAB Fall Meeting

CVDI 2019 IAB Spring Meeting

CVDI Calendar

Project - Benefits to IAB

- Nearest neighbor search is a very standard operation ini information retrieval, clustering and visualization. Being able to scale up the methods opens up new possiblities. The actual business cases with interested IAB members will be planned later.

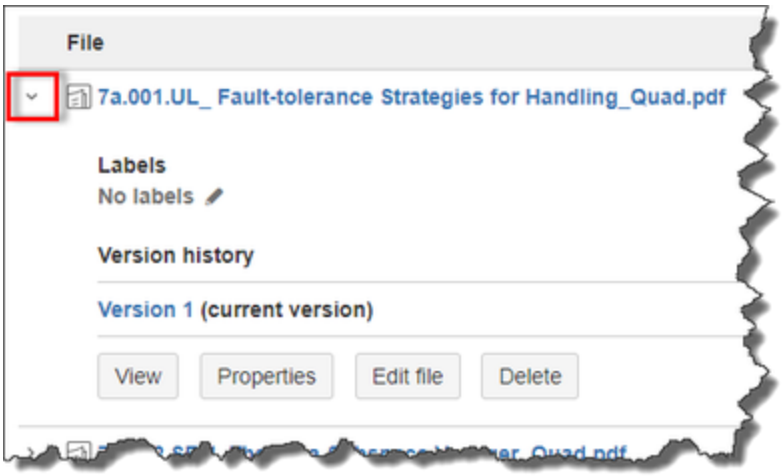
Project - Presentation Video (Spring 2018)

Video Link (8:47 minutes)

Project - Documents

For viewing/editing options, please click left arrow next to document name.

You will see different options depending on your access level.



File	Modified
7a.029.TUT_Very Fast Nearest Neighbor_Retrieval_Quad_2017 Fall Meeting.pdf	Jan 05, 2018 by Sally Johnson
7a.029.TUT Executive Summary.pdf	Feb 22, 2018 by petri.myllymaki@helsinki.fi
7a.029.TUT-Very-Fast-NN-Presentation-with-audio.ppsx	Feb 22, 2018 by Moncef Gabbouj
7a.029.TUT-Very-Fast-NN-Presentation-with-audio.mp4	Feb 22, 2018 by Moncef Gabbouj
7a.029.TUT_Quad Chart_2018 Spring Meeting.pdf	Mar 15, 2018 by Sally Johnson
7a.029.TUT_2018 Fall Meeting Poster.pptx	Nov 09, 2018 by Sally Johnson
7a.029.TUT_Mid-Year Report.pdf	Jan 04, 2019 by Sally Johnson

Drag and drop to upload or browse for files

CVDI Leadership (All Sites)
CVDI Marketing Materials
CVDI Reports & Document Library
CVDI SITE (Drexel University)
CVDI SITE (Stony Brook University)
CVDI SITE (Tampere University)
CVDI SITE (University of Louisiana at Lafayette)
CVDI SITE (University of North Carolina at Charlotte)
CVDI SITE (University of Virginia)
IAB - Industry Advisory Board
Year 6 - Funded Projects (7/1/17 - 6/30/18)
Year 7 - Funded Projects (7/1/18 - 6/30/19)
Year 8 - Proposed Projects

Download All

Project - Comments