5 Research projects*

BioInfStat -   Statistics for bioinformatics
ClimateInsure -   Climate Change and Insurance Industry
ComplexClin -   Statistics for complex design of clinical studies
ComplexDepend -   Statistics for complex stochastic dependence
CustomerLife -   Statistics for modelling customer life in an insurance company
Elprice -   Electricity price sensitivity
FindOil -   Statistics for oil and gas exploration
Genestat -   Statistics for genomic research
Infect -   Modelling spread of infectious diseases in fish farming
StatMarine -   Statistics for management of Norwegian marine resources
TotalRisk -   Statistics for modelling the risk of financial institutions

BioInfStat
Statistics for bioinformatics

Innovation area: Health

Key Innovator: Marit Holden

Partners: OUS, Biomolex, PubGene, UiO, NR

Research staff scientists: Marit Holden (principal investigator, NR), Eivind Hovig (principal investigator, OUS), Line Eikvil (principal investigator, NR), Ingrid Glad (UiO), Knut Liestøl (UiO), Håvard Hauge (Biomolex), Fang Liu (PubGene), Tor-Kristian Jenssen (PubGene), Daniel Rustad (PubGene), Harald Kirkerød (PubGene), Arnoldo Frigessi (UiO), Hugo Hammer (HiOA)

Additional reference group: Eirik Næss-Ulseth (PubGene), Torbjørn Rognes (Sencel)

International contacts and collaborators: CAS (the Laboratory of Complex Systems and Intelligence Science in the Institute of Automation, Chinese Academy of Sciences in Beijing, China), Bo Jansson (BioInvent International AB, Lund,), Sven-Erik Strand (Univ. of Lund)

Scope:

The aim is to apply, adapt and develop new statistical methods to enhance already existing biotechnologies to make them more competitive.

Results in 2012 and plans:

1. Tools for biomedical search
The overall goal for the activities related to PubGene is to provide users with text mining tools that enable them to easily get an overview of the relevance of documents returned from a biomedical search and efficiently single out the useful information. Through the project such solutions have been developed and integrated into the CoreMine system, and as a result users

---

* Take contact with a principle investigator for more information on our projects, or if you are interested in joining a project.

can now make a keyword based search in PubMed, cluster the returned results and receive clusters accompanied by textual labels and summaries describing the contents. This is also integrated with the network concept of CoreMine, enabling visualization of cluster relationships in a network.

Through our work we have seen that what is the best clustering, cluster labels or cluster summaries very often depends on who you are asking. Different users will have different needs and interests, e.g. a biomedical researcher may see things differently from a GP. Therefore we have introduced the concept of multi-focus to give users the possibility to get an overview of the contents from multiple viewpoints. In the remainder of the SFI-period we plan to develop this concept further into a solution that can be demonstrated in CoreMine, and to write a paper describing this new concept. In addition, we plan for refinement of the presentation and visualization of the results from the new tools that have already been integrated in CoreMine.

2. The imaging process of the Biomolex detector
Biomolex has technology for real-time imaging of radioactive emissions. Some mechanisms in the imaging process are not sufficiently well understood and need to be further investigated. Biomolex will therefore produce data for phantoms of selected isotopes. We will then analyze these data to obtain a better description of how both the selected and other isotopes behave in the Biomolex detector.

Papers:

Eikvil, Line; Holden, Marit and Jenssen, Tor-Kristian: Organizing biomedical literature abstracts through clustering, cluster labelling and cluster summarization. In preparation 2012.

Technical reports:

Eikvil, Line; Holden, Marit: Word sense disambiguation and co-occurrence ranking, Applied to PubMed abstracts. NR-Note, SAMBA/46/09, 2009.

Eikvil, Line; Holden, Marit: Efficient search for biomedical information. NR-Note, SAMBA/54/11, 2011.

Holden, Marit; Glad, Ingrid; Hauge, Håvard; Liestøl, Knut: Image restoration and analysis of biomolecular radioactivity images, NR-Note, SAMBA/27/08, 2008.

Holden, M; Glad, IK; Hauge, HH; Hovig, E; Liestøl, K. Image restoration and analysis of biomolecular radioactivity images II. NR-Note, SAMBA/11/12, 2012.

PhD defence of Hayat Mohammed

(sfi)²