Job Description – Data Scientist

Location: **Oxford, UK or Hørsholm, Denmark**

Type: **Full-time, permanent**

Job Ref Nr: **SYN-DS-01/20**

Reporting To: **Haja Kadarmideen, CTO**

Synomics aims to transform the efficiency of the agri-food industry by improving animal and crop production efficiency, and the economics and sustainability of the entire production value chain for multiple species. Our platform is a revolutionary advance in the analysis of high-throughput genotypic, phenotypic, and environmental data to accurately measure and analyse protein production efficiency, health, accelerate genetic progress, and deliver more sustainable protein production from animals and plants. We're a small but rapidly growing AI enabled genomics, breeding and precision agriculture company with global ambition and partnerships with some of the largest international livestock breeding companies and research institutes.

We're located near Oxford, UK and in Hørsholm (near Copenhagen), Denmark and were founded by a diverse team of highly experienced animal production & health, genomics, breeding, bioinformatics, AI and computer science leaders.

We have a fantastic opportunity for a Data Scientist to join our team, reporting to our CTO, Professor Haja Kadarmideen. We are looking for someone who is comfortable working with large and complex datasets/databases and who has significant experience building data pipelines at scale for our animal/plant genomics and breeding company. The ideal candidate will have a relevant educational background (computer science, applied mathematics and/or statistics) and would ideally be comfortable working in the agricultural industry. They will help other team members to identify and commercialise novel correlations within large and complex life science and animal/plant phenotypic, genomic and bioinformatic datasets.

You will furthermore be involved in close cooperation with quantitative geneticists, bioinformaticians, biostatisticians, computational biologists and software engineers assisting with programming efforts, genomic and bioinformatics software tool development, data design and workflow management.

This is an exciting opportunity to help shape the direction of the business by accelerating technology change, improve animal welfare and reduce carbon emissions across a globally significant industry. The ideal candidate will have a keen interest in creating insight and value from complex data, and a passion for asking and answering questions that help us solve real-world global challenges and build the best products in commercial environment.

Responsibilities

- Building data pipelines at scale using a range of tools and technologies and identifying opportunities for future innovation of genomics, multi-omics and phenotype data exploration
- Designing, coding, testing and documenting data analysis scripts using agreed standards and tools.
- The QA/QC, preparation and analysis and visualisation of datasets using tools such as R and/or Python.

- providing data science expertise and guidance to quantitative geneticists, bioinformaticians, biostatisticians, computational biologists and software engineers in the selection of methods for specific projects and products
- contributing to the identification and/or development of new analytical methods for creating insight and value from complex datasets
- communication of technical and analytical subjects between different specialities within the business, and with partners and customers
- working with the Executive Team to design and develop product roadmaps for new analytical methods
- ensuring best practice and methodologies use used for the QC of data and results
- ensuring IP opportunities arising from research are identified
- acting as an expert internally and externally in relation to the analysis of data using company products
- writing-up and publishing new methods and results
- where appropriate, to provide mentorship and advise in statistics and data science to the wider company

This job description is not exhaustive and will be subject to periodic review. It may be amended to meet the changing needs of the business

Key Requirements

- MS/MSc. or Ph.D. in computer science, applied mathematics or statistics or a relevant technical field, or 3+ years' experience in a relevant role
- In-depth knowledge and experience with R and preferably also in Python and ability to use them in data processing pipelines and algorithm application/development
- Experience in use of machine learning/AI methods for big data integration, analysis and visualisation,
- Experience in applied statistics (ideally biostatistics), modelling, quality assessment/control, data exploration, and clustering techniques
- Database skills, preferably relational databases (MS Access/SQL) in design, maintenance, query and performance
- Ability to develop software / algorithms for processing and statistical analyses in proof of concept and commercial projects
- Ability to solve analytical problems using quantitative approaches
- Experience developing (in addition to applying) computational tools and algorithms in the fields of bioinformatics, machine learning, statistics or other related field
- Experience in parallel programming models, Linux or HPC clusters, and/or CUDA computing would be a plus

Key Attributes

- Good communication and presentation skills including in English (min. of equivalent of CEFR level C1)
- A strong passion for empirical research and for answering hard questions with data
- A willingness to dive deep, experiment rapidly and get things done
- Ability to work in a diverse team of geneticists, data scientists, biostatisticians, bioinformaticians, breeding specialist, software / IT specialists and business support across multiple domains/species
- Confidence to make suggestions and influence change
- Ability to work well to deadlines, while paying attention to detail and accuracy
- Flexible approach to working
- A high degree of pragmatism in building the simplest thing that does the job; we want to use technology to solve real-world problems expediently
- Personal integrity and discretion

What We Offer

- Competitive salary with performance related bonuses
- Company pension scheme
- 25 days annual leave rising with long service
- Excellent career development opportunities in a fast-growing international organisation

Application Process

- Please submit cover letter, your CV, and publication list (if relevant) to jobs@synomics.ai
- Applications close on 27 April 2020
- This position is available for immediate start after the interviews

Please mention the job reference number SYN-DS-01/20 in the cover letter and subject line of your email when applying.

Synomics is a collaboration between Wheatsheaf Group Ltd and PrecisionLife Ltd. Wheatsheaf Group directly operates, invests in and helps to develop businesses in the food and agriculture sectors. PrecisionLife is an AI enabled precision medicine company, which is ISO27001 accredited, committed to equal opportunities and operates a carbon neutral business. To learn more, visit:

http://www.synomics.ai/, https://www.wheatsheafgroup.com/ and https://precisionlife.com/.

We follow the EU right to work employment checks and you will therefore need to satisfy basic eligibility criteria/certain conditions of employment (e.g. nationality rules/right to work in EU / UK); and provide appropriate original documentation to verify ID, nationality, employment and/or academic history.

Synomics is fully committed to equal opportunities.

Here's a link to our privacy policy- http://www.synomics.ai/privacy-and-cookies-policy/

In this policy, you will find information about our compliance with GDPR (data protection law). You can find how to send us a request to let you access your data that we have collected, request us to delete your data, correct any inaccuracies or restrict our processing of your data. You have the right to lodge a complaint about the way we handle your data with supervisory authority or you can contact our DPO at GDRP@synomics.ai for more information or concerns. We process this data for recruitment purposes only. We store all candidate data in secured file with restricted access, and we will not share it with anyone else.