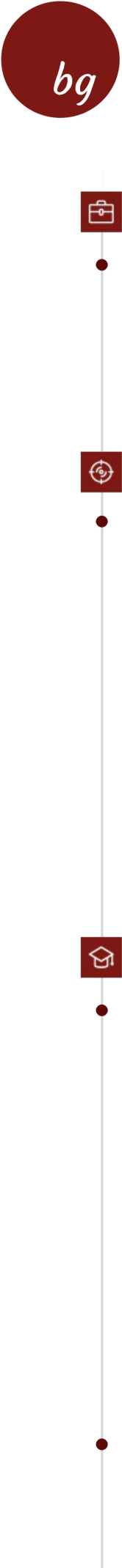
**Bishwa K. Giri**

**Data Scientist, Python Developer**

**Contact :** [kirannbishwa01@gmail.com](mailto:kirannbishwa01@gmail.com) ,

[bishwa.giri@wellsfargo.com](mailto:bishwa.giri@wellsfargo.com),336-3929572



**Professional Summary**

* 7+ years of experience in application development and implementation, application optimization and big data analyses.
* Advanced Python and R programming skills, bash and shell scripting skill with 7+ years of experience.
* Regular contributor to the open source application development and implementation.
* Experienced in DevOps and Agile project methods for achieving faced paced project goals.
* Experienced in collaborative project management and reporting using JIRA and Confluence.
* Experienced in Information Security methods and tools with knowledge on Tanium, Splunk, NDEX, Threat Connect, Elastic Search, etc. for analyzing machine data to generate threat intelligence.
* Experienced in handling large data sets using tools like pandas, numpy, dask etc.
* Experienced in working with larger than memory databases using Dask.
* Good knowledge of data architecture used in NumPy, Pandas and Dask - e.g : arrays, dataframe, hdf5 format, etc.
* Extensive experience in data filtration, wrangling, transformation and creating data analyses pipelines.
* Extensive knowledge in algorithm design, implementation and application debugging and optimization.
* Experienced in application and package version control and management using tools like conda, virtual environment, pyinstaller, docker etc.
* Experienced in TDD (test driven development) for efficient application development.
* Experienced in efficient and (CI/CD) continuous integration and deployment using tools like Git, GitHub Actions etc.
* Experienced in data visualization using python and R libraries – matplotlib, ggplot2, plotly, plot etc.
* Experienced in optimizing application and data analyses pipelines via multiprocessing, multithreading, memory management, run time reduction.
* Extensive experience with all major OS (Window, Unix, Mac).
* Experienced in TCP/IP, HTTP, FTP and several other internet and network protocols.
* Experienced in running jobs on HPC (high performance computation) cluster and large databases using SSH, TLS, SSL protocol.
* Experienced in handling SQL, NoSQL and JSON databases.
* Highly skilled in interdisciplinary framework with integrative knowledge in bio-informatics, information security, statistics, big database use and management, scripting and application development.
* Advanced knowledge of several bio-informatics tools – regularly analyzed large sequence datasets (+500 GBs) of sequence reads using command line applications like BWA, GATK pipeline, Picard, phaser, rnaSTAR, g2gtools, samtools, bowtie, tophat, bcftools, vcftools, emase, IGB, IGV, GenomeBrowser, QCtools, BLAST, BOWTIE, MUSCLE etc.
* Regularly analyzed alignment and variants information from genomic and transcriptomic databases – SAM/BAM files, VCF, GTF, GFF, fasta, fastq, RefSeq, HapRefPanel.
* Contributed independently to the algorithm, design and development of several large scale application – developed several full length open source bioinformatics application and data pipelines for analyzing big databases from genomics and RNAseq technologies at UNC Greensboro. Also, forked and optimized other public repositories involved in bio-informatics data analyses.

**Applications hosted on GitHub @** [**https://github.com/everestial**](https://github.com/everestial).

* + - * + **Developed data processing pipeline** [***aseCADG***](https://github.com/everestial/ASE-CADG) **for analyzing ASE (allele specific expression) from RNAseq reads of the F1 hybrids.**
        + **Developed haplotype phasing application using first order markov chains -**[***phase-Extender***](https://github.com/everestial/phase-Extender) ***and*** [***phase-Stitcher***](https://github.com/everestial/pHASE-Stitcher) ***that forms a part of***  [***phaseIT***](https://github.com/everestial/PhaseIT/tree/master) ***pipeline.***
        + **Developed python application** [***VCF-Simplify***](https://github.com/everestial/VCF-Simplify)**for simplifying and building VCF files for easy VCF data manipulation.**
* Experienced in google cloud genomics data analyses and data processing pipeline setup.

**Tech Skills**

* **Office Productivity :** Microsoft Office Package (Word, Excel, PowerPoint), LibreOffice, Google docs.
* **Operating Systems :** Windows 98, 2000 Server, XP, Vista, 7 and UNIX/Linux, Ubuntu.
* **Security and log management tools**

Tanium, Splunk, Threat Connect, Elastic Search, NDEX (network data explorer).

* **BioConductor packages**

DESeq2, edgeR, karyoploteR, haplotypeR, clusterprofiler, DOSE, etc.

* **R packages and management tools**

Bioconductor, ggplot2, plotly, tidyr, dplyr, shinyr, enrichGO etc.

* **Python Packages**

pandas, dask, pytest, pyspark, numpy, Sqlite, shlex, cyvcf2, matplotlib, Django, TkInter, keras, multiprocessing, scrapy, kivy, BeautifulSoup, etc.

* **IDE (integrated development environment)**

Rstudio, Rcmdr, JupyterLab, Pycharm, Spyder, Anaconda navigator.

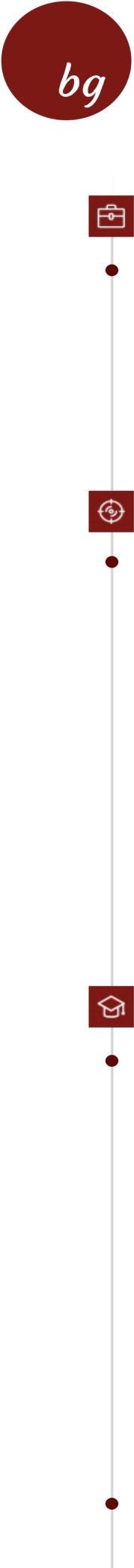
* **Programming and Scripting tools/environments :**

bash, sed, awk, grep, Python, R, Go lang, JavaScript, PowerShell

* **Statistics:**
* Advance knowledge of both frequentist and Bayesian statistics.
* Experienced in statistical data analyses using regression, ANOVA, markov chain, MCMC (markov chain monte carlo) models, hidden markov model, etc.
* **HPC (high perfomance computation) :** Experienced working in HPC system with job submission, data query and database management.
* **IDE for application DevOps:** Experience with version control/integration/tracking tools like Git, Subversion, Jenkins.

**Soft Skills**

* Excellent interpersonal and time management skills and motivated by challenges and ability to perform well under pressure and multi-task effectively.
* Excellent communication skills in both verbal and professional writing. Fluent in languages : English, Hindi, Nepali.
* Ability to work in a team to achieve team goals.
* **Willing to learn more and open to the ideas of application and methods developments for advancement of STEM (Science, Technology, Engineering and Mathematics), interdisciplinary projects , open source projects and research.**
* **Willing to actively engage in the use of new generation of technologies like machine learning, deep learning, cloud computing etc.**

 **Experience**

**Information Security Engineer 3 for Wells Fargo via. Apex Systems (Sept, 2018- October 2019)**

* Involved in the continued maintenance and enhancement of the python application/packages with new security requirements. These packages form a part of security tool implemented in Tanium Action.
* Involved in preparing test and implementation documentation of the developed python packages.
* Developed unix sensor for Tanium Interact that helped to communicate with the end points and gather results for specific questions.

**Information Security Engineer 3 for Wells Fargo via. Apex Systems (Sept, 2018- October 2019)**

* Developed python application/packages to parse security patch logs applied to Wells Fargo devices to validate security updates.
* Developed cross platform scripts in python to transform data between different formats (delimited text, excel, JSON etc.).
* Extracted machine data from Splunk indexes to monitor, research and provide cyber threat intelligence.
* Involved in the development of MCMC (markov chain monte carlo) statistical models to generate quantitative risk assessment.
* Involved in the development of network visualization of network/hosts/ip to provide threat intelligence monitoring.
* Used tanium, ThreatConnect, NDEX to monitor and further refine security data logs to reduce false positives.

**Graduate Assistant (7+ years) for UNC Greensboro, Biology Department (2011-2018)**

* Taught General Biology Lab courses (Bio 105, Bio 111 & Bio 112) and upper level Bio 392 (Genetics) Lab.
* Trained undergraduate researchers in several biology, bio-informatics and data analyses research projects.

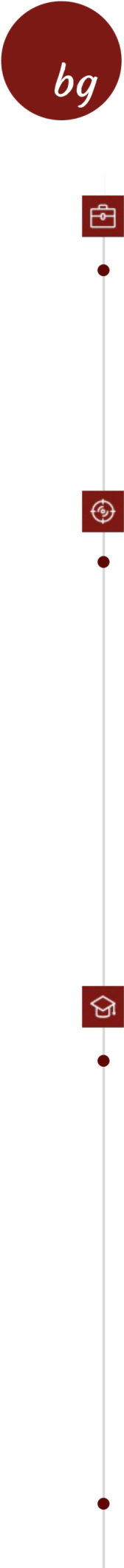
**PhD researcher and bioinformatics data analyst (7+ years) in the lab of Dr. David Remington, UNC Greensboro**

* Researched evolution of life history tradeoffs in [***Arabidopsis lyrata***](https://plants.usda.gov/core/profile?symbol=arly2).
* Actively involved in genomic DNA extraction, RNA extraction, next generation sequence reads library preparation.
* Generated hypothesis and predictions about genomics and RNAseq databases to associate gene to phenotype (lifehistory traits) in [***Arabidopsis lyrata***](https://plants.usda.gov/core/profile?symbol=arly2).
* Developed statistical methods for test of hypothesis and interpretation of the results.
* Analyzed +500 GBs of next generation sequence data (genome resequence and RNAseq reads) using softwares like BWA, rnaSTAR, phaser, GATK, Picard, etc.
* Regularly prepared data analyses jobs and submitted it to the HPC cluster (Henry on NC state university) using ssh protocol.
* Created databases of aligned BAM files, variant call files (VCF), quality metrics, statistics of the observed variant from the aligned sequence reads.
* Developed algorithm using first order markov chains for classifying haplotype blocks for haplotype phase extension.
* Developed ASE (allele specific expression) pipeline [***aseCADG***](https://github.com/everestial/ASE-CADG) [for the test of ASE (allele specific expression) in F1 hybrids.](https://github.com/everestial/ASE-CADG)
* Developed full length python based haplotype phasing application – [***phase-Stitcher***](https://github.com/everestial/pHASE-Stitcher)and[***phase-Extender***](https://github.com/everestial/phase-Extender)and established haplotype phasing pipeline [***phaseIT***](https://github.com/everestial/PhaseIT/tree/master).
* Created databases of genome wide phased haplotype for all the sequenced samples using the applications [***phase-Stitcher***](https://github.com/everestial/pHASE-Stitcher)and[***phase-Extender***](https://github.com/everestial/phase-Extender).
* Identify problem trends pro-actively.
* Maintain information confidentiality and security.
* Analyze, troubleshoot and report any issues on the file/application server systems.
* Analyze the output files and verify that it confirms technical requirements.
* Involved in further curation of the output database, data management, query, and statistical analyses of the refined dataset to test the hypothesis.
* Developed statistical models for count data analyses to resolve ASE (allele specific expression) patterns in RNAseq data from F1 hybrids and established the pipeline [***aseCADG***](https://github.com/everestial/ASE-CADG) .

**Undergraduate Researcher, Institute of Forestry, Pokhara, Nepal**

* Wildlife habitat mapping and ecological modeling using ArcGIS, EdrasImagine.
* Collected database on species distribution using field survey, catch and release technique.
* Generated hypothesis and tested data using statistical tools SPSS, R.

**Thesis: “**Habitat suitability Mapping and Species Identification of Chiroptera: A case Study from Kaski District, Nepal.” Thesis submitted to Tribhuvan University, Institute of Forestry, Pokhara Campus, Pokhara.

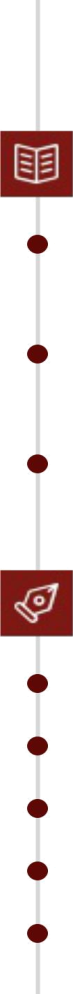
 **Education**

**PhD (Environmental Health Science with focus on Bio-Informatics), 2011- Current (Expected Graduation, Dec 2020)**

University of North Carolina at Greensboro, NC

**B.Sc Forestry, 2005 – 2009**

Institute of Forestry, Pokhara, Nepal

**Projects/ Publications (on track)**

* Haplotype phasing in heterogenous genome and F1 hybrids using [***phase-Extender***](https://github.com/everestial/phase-Extender) ***and*** [***phase-Stitcher***](https://github.com/everestial/pHASE-Stitcher) ***.***
* Analyses of allele specific expression in *A. lyrata* F1 hybrids using [***aseCADG***](https://github.com/everestial/ASE-CADG) - pipeline for competitive alignment of sequence reads on a diploid genome.
* Auxin inhibition affects lateral shoot development and alters life history traits in *A. lyrata.*

**Personal Sites:**

**GitHub :** [*https://github.com/everestial*](https://github.com/everestial)

**LinkedIn :** [*https://www.linkedin.com/in/bishwa-kiran-84628728/*](https://www.linkedin.com/in/bishwa-kiran-84628728/)

**Wordpress:** [*https://wordpress.com/post/everestialblog.wordpress.com/21*](https://wordpress.com/post/everestialblog.wordpress.com/21)

**StackOverflow:**

[*https://stackoverflow.com/users/story/6346698*](https://stackoverflow.com/users/story/6346698)