NSDC-HerWILL: Data Forge 2023 Report

Picture of Instructor

Introduction of Instructor

Raima Islam is a recent Computer Science graduate from BRAC University. She is currently a research fellow under the Fatima Fellowship, conducting research on computational biology and machine learning at Harvard University and Stanford University. Additionally, she is an AI Fellow of Italy's Pi School of AI, exploring how to use Large Language Models to solve business-oriented challenges.

Data Forge 2023

HerWILL - Women Inspired in Life and Leadership - organized their signature Data Science program: Data Forge 2023 consisting of workshops and a contest across the globe in December 2023 in collaboration with the National Student Data Corps (NSDC) HerWILL Chapter. The program consisted of a series of custom-made workshops to give an in-depth overview of Data Science to the participants and an intense 72-hour Datathon.

Story in numbers

540	51	95	28	5	3
Workshop Participants	Contest Participants	Participating Universities	Cities	Worldwide Mentors	Winning Contestants

Description

In this unique program, participants from different corners of the world participated in extensive and dynamic workshops encompassing a full pipeline data science project along with a contest on a challenging dataset to solve a real world problem provided by Kaggle.

Around 6 top Data Scientists from four countries gathered to instruct, supervise, and mentor the students for the program.

The top three winners collectively received certificates from the NSDC HerWILL Chapter, along with Grow with Google scholarships, enabling them to become the next generation of leading Data Scientists.

Pre-Workshop Expectations:

- Gain foundational knowledge and skills in data science, including coding, machine learning, and analysis.
- Understand practical applications of data science in real-world scenarios and industries.
- Enhance coding skills through hands-on experience in data science projects.
- Explore the intersection of data science with various fields like mechanical projects, economics, and chemistry.
- Obtain insights from industry specialists in the data science community.
- Develop a career roadmap, considering the job market's demand for data science skills.
- Participate in workshops and contests applying theoretical knowledge to practical problems.
- Network with experts and professionals in data science for potential collaborations and mentorship.
- Gain exposure to tools, techniques, and methodologies used in data science for effective data analysis.
- Receive certificates or achievements to enhance resumes, improve employability, and open doors to scholarship opportunities for further education.

The Proud Winners

- Mashiyat Mahjabin Prapty - BUET
- Muniza Alam - IUT
- Deen Monhammad - SUST

The Winners' Experience:

Mashiyat Mahjabin Prapty: I am Mashiyat Mahjabin Prapty, currently a final year student in Bangladesh University of Engineering and Technology. I am passionate about Data Science and my thesis work is related to the application of Machine Learning in the medical field. In my leisure time I read books and dream about travelling the whole world someday.
Experience: NLP was a new field for me before this datathon. I learnt a lot in the workshops and tried to apply the newly learnt knowledge in the datathon. Thanks to HerWill and NSDC for giving me this opportunity. I enjoyed participating very much.

Muniza Alam: I'm Muniza Alam, rounding off my EEE degree at IUT. Tech has always been my playground. Currently navigating through Capture the Flag competitions and diving into the intriguing universe of machine learning. In my downtime, you'll find me unwinding with a book, exploring worlds beyond circuits and algorithms.
Experience: Participating in the Data Forge 2023 contest was an enlightening journey into the realm of data science. As an EEE undergrad with limited coursework related to data science, this 3 day long workshop & contest strengthened my base & clarified many queries. The workshop covered crucial aspects of Natural Language Processing (NLP), decoding data and classification using supervised learning. The hands-on experience during class, as well as the contest, allowed me to craft unique projects for my portfolio, enhancing both my practical skills and theoretical understanding. The instructor was super helpful and well-informed regarding errors & troubleshooting. This is a comprehensive package that adds immense value to my learning experience.

Deen Mohammad: I am Deen Mohammad, a fresh graduate in mathematics from Shahjalal University of Science and Technology, Bangladesh. Currently, I am serving as a data analyst for a private company. Fascinated by the potential of predicting the future with high confidence, I became passionate about data science. My goal is to contribute to any innovation in this field that will be impactful to humanity.
Experience: I've had the incredible privilege of participating in Data Forge 2023 by HerWill, an experience that I will remember for a long time. After three days of rigorous training sessions on natural language processing, this competition was a chance to buzz with creativity. The workshops were informative and gave hands-on data science learning experiences to all participants. I am delighted to win the second runner up position at the competition and to be part of this supportive community with a great passion for data science. I greatly appreciate the organizers dedication and initiative to provide a wonderful platform, especially for women around the world, to learn data science and transform their careers.

Key findings

- **Technical Skill Improvement:** Events such as HerWILL's Datathon improve Data Science skills very fast. With 3 workshops, HerWILL's event was able to improve more than half of its participants' Data Science technical skill, where their average score improved from a pre-event 2 to a post-event 4 on a 1-5 scale.
- **Potential Enthusiasm:** Data Science is a necessary tool in the fast advancing world and we sensed great enthusiasm among our participants to learn this thing. Participants from Bangladesh outweighed participants from all other countries, which shows high acceptance of our programs in this region.
- **Necessity of More Such Programs:** As we proceeded with our workshops towards the 72-hour Datathon, we collected feedback from our participants. Continuation of such programs are requested by more than 50% of them.

Top Participating Universities

- University of Chittagong
- Bangladesh University of Engineering and Technology (BUET)
- BRAC University
- University of Asia Pacific (UAP)
- Shahjalal University of Science & Technology (SUST)
- Mount Holyoke College
- Temple University
- University of Massachusetts Amherst
- Islamic University of Technology (IUT)

Top Performers Came From

- Bangladesh University of Engineering and Technology (BUET)
- Islamic University of Technology (IUT)
- Shahjalal University of Science & Technology (SUST)

The Program Model

HerWill's Data Science Workshop 2023 was made dynamic with the objective of meeting the evolving needs of our industry and academic partners. Our model included:

- Three 1.5-hour online workshops with live instructors.
- Instructors that are experts who provide real-life examples to bring content relevance and thus make Data Science less intimidating.
- 100% of the program was run virtually crossing geographic boundaries and time zone issues.
- Custom-built project-based workshops were held to give an overview of Data Science. The courses went from basic to intermediate levels in depth subject areas to ensure success for all contestants.

Dataset Came From Kaggle to:

- Detect fake news using Natural Language Processing (NLP) and Machine Learning (ML).
- It includes instruction on visualizations and deploying models for real-world applications.

The Global Leaders

- Sai Gattupalli - UMass Amherst
- Scott Blender - Temple University
- Irfan Nafiz - SUST
- Mubin Khan - SUST
- Ahmad Mustafa Anis - ML Engineer at RedBuffer
- National Student Data Corps HQ team

Workshops Led by the HerWILL Team

- Decoding Data: Insights and Impact
- Introduction to NLP Processing
- Classification using Supervised Learning

Post Workshop Feedback:

- Gratitude for the free sessions and praise for the instructor, citing them as amazing.
- Participants expressed satisfaction with the informativeness and educational value of the workshop.
- Positive comments about the session; a desire for more similar workshops in the future.
- Desire for additional resources and tips to stand out in job searches.
- Appreciation for the opportunity and interest in more workshops covering various data science topics.

HerWILL workforce development overview

HerWILL has developed a program that brings together project-based, hands-on exercises delivered by experts in various fields of Data Science, industry leaders, and academic researchers to launch and grow careers in Data Science and AI. The multi-year, multi-faceted program built a community of Data enthusiasts that will be future leaders and may be matched with hard-to-reach opportunities.

Why data science

Women represent only 26.7% of the tech workforce. On the flip side, Data Science jobs are in the top five highest paying opportunities with the greatest demands for talents. Our goal is to bridge this gap in high demand and low supply of skilled talents through Data Science education, benefiting women with better economic opportunities and organizations with better talents. HerWILL's program will help talented women get recognized, trained, and certified to play a bigger role in the global tech future.