The Art of Bioinformatics Learning in Our Arabic World

Alsamman M. Alsamman *[1]

[1] Molecular Genetics and Genome Mapping Lab. , Agricultural Genetic Engineering Research Institute, Agricultural Research Center, Egypt.

* To whom correspondence should be addressed: smahmoud@ageri.sci.eg

Abstract

Bioinformatics became a significant field in life sciences that, draws a number of researchers and extends into a wide range of biological disciplines. Rendering bioinformatics analysis techniques are the most desirable skills in a variety of scholarship programs and academic positions. Teaching bioinformatics is very challenging since it is a multidisciplinary field, where most of the undergraduate programs in colleges provide only one area required for bioinformatics. Besides the regular education system, few bioinformatics training courses are offered and less are affordable to fresh graduates in countries most of which are categorized as developing countries. The high cost of learning, confusing education systems, and the complexity of bioinformatics science has made it very difficult to be taught and more challenging to be studied in Arab countries. This review provides possible solutions to most of these issues and offers the best practice to guide future Arab bioinformaticians to learn bioinformatics in a way that fits our social, financial and academic circumstances. Moreover, it discusses the key aspects that a bioinformatician needs to be aware of and the basic knowledge that must be gained. On the other side, it will illustrate how to start learning, to address some of these challenges and how to deal with some of the related social issues.

Keywords: bioinformatics, Arab, programming, education, Linux, biological data.

Introduction

Bioinformatics analysis techniques are the most desirable skills in a wide range of scholarship programs and academic positions. As the study of bioinformatics and computational biology grows and evolves, it is essential to quantify the factors that contribute to the development of professionals in this field (1). Bioinformatics is an interdisciplinary research field where computational resources and techniques are used to interpret biological data through mathematical and statistical approaches.

Data Availability Statement: All relevant data are within the paper

The rapid acceleration in computing power and memory storage capacity have given rise to a new golden age in the biological data analysis (2). There is almost no biological sector in which bioinformatics has not yet been incorporated. Its techniques are used in microbiology to examine microbial diversity and species occurrence by identifying and quantifying the association of microbial communities among different biological samples (3). Throughout the pharmaceutical industry, bioinformatics offers analytical tools that can enhance drug target identification, drug candidate monitoring and drug optimization. In particular, it promotes the recognition of side effects and forecasts drug resistance (4).

Recently, bioinformatics has proposed the idea of personalized medicine, where treatments will be adapted to the unique genotype of patients. Integrating the vast genetic information provided by the Genome Wide Association Studies (GWASs) is a valuable resource for mapping genetic traits with drug reactions and phenotypes, allowing individual characteristics of each patient to be monitored and their susceptibility to certain diseases to be considered (5).

Last but not the least, the integration of machine learning in bioinformatics analytical methods has opened a new era in which sample data and past experience could be used to improve the output criterion in computational algorithms, the optimized criteria could be the rating of reliability provided by the statistical model and the significance of the performance (6). Machine learning has therefore enabled what appears to be a breakthrough in biological research, where computer programs can solve complex biological problems efficiently and effectively. Through cancer diagnosis (7), neurooncology imaging (8) and drug design (9) in medicine and plant physiology (10), forecasting crop yields (11) and livestock production in agriculture (12), machine learning has been used to solve both challenging and basic tasks.

The rapid development of life sciences and information technology requires the continuous development of bioinformatics learning programs in order to sustain their significance (1). The quality of the education systems in the Arab countries has gradually improved over the last 30 years in the Arab world. More work has been done by non-governmental and government agencies and organizations to improve the educational opportunities for Arab students, and to promote their independence and integration into their societies and to prepare them for future careers, taking into account current trends in the labor market (13). Few leading bioinformatics programs in Arab countries have been established over the last decade in Egypt, Lebanon, KSA, UAE, and Oman. Some of these programs have launched throughout the biotechnology and genetics departments of computer science, science and agriculture colleges in both special and governmental universities (14). So far, in Arab countries, neither colleges nor departments have been specifically founded to teach bioinformatics.

As multidisciplinary science, bioinformatics is difficult to teach, most of the programs are designed for undergraduate students in colleges that provide only one field required for bioinformatics. These programs address these issues by borrowing a few courses from other colleges, for example, when bioinformatics programs are embraced by computer science colleges, students are expected to pursue a few courses in biology in science or agriculture colleges. These courses could take place in the second or last years, and could be acquired in the worst situations a few weeks before graduation. This system creates a gap in the experience of the participant, a hole that must be filled in the next few years of his life as a researcher or even as an employee.

Beyond the regular education program, few bioinformatics training courses are offered and less are affordable to fresh graduates in countries most of which are classified as developing countries. The high cost of training, confusing education systems and the complexity of bioinformatics science has made it very difficult to teach and more difficult to study in Arab countries. Inducing bioinformatics education systems to move slowly toward improving the quality of life of Arab residents and addressing crucial food and drug issues.

This review is written in order to offer possible solutions to most of these issues and to guide future Arab bioinformaticians throughout the best methods to study bioinformatics in a way that fits our social, financial and academic circumstances. It also addresses the key aspects that the bioinformatician needs to know about, such as how to begin learning the basic knowledge that needs to be learned, and how to overcome some of the social issues faced by young bioinformaticians.

How to get started in bioinformatics

Bioinformatics, as already stated, comprises of three different fields: genetics, computer science and mathematics. The purpose of these three domains is important for the organization, comprehension and interpretation of the different biological information that bioinformatician handles on a daily basis. The question of studying these three areas is very difficult and few places of education provide this knowledge in one location.

You've certainly learned the fundamentals of one or two of these areas as biology, computer science or mathematics graduate, you only need one or two remaining knowledge to become a bioinformatician. As you begin to learn, my tips are:
1) Learn the basics, and then you will know where to go next.

2) This is about knowledge, not certification; certification will not indicate that an individual has adequate knowledge of the subject area. Arab bioinformatics students are more worried about the credential than they are about the skills they need to acquire.

3) Take what you need through free learning; this is your university and your home, attend those classes you need to learn, even if it is not your college. Many students overlook the fact that most university lectures are not closed to free learners as long as they do not need a credential.

4) Take a basic course; most of the basic courses offered in Arabic countries cover weak bioinformatics skills such as primer design, sequence alignment, NCBI software and gene annotation using online tools. You may take one basic course if you don't understand the software, as long as you don't repeat courses with the same content.

5) Ask as much as you can; there are many public forums for bioinformaticians that offer gatherings across all disciplines such as biology, mathematics, and computer science, where you can ask any question you want.

6) Engage the scientific society as a free bioinformatician; I know it's hard for some Arab students to work after graduation for 6-12 months without any kind of support. But it is about joining the research community , learning about the issues they face and seeking out how to manage it as a bioinformatician. The advantage of this would be that, if they don't pay they can't control you. You can choose to work on any type of data without any restrictions, the ability to attend any class outside the workplace, and if there is any other work opportunity you can apply. In fact, you should have a certificate of work experience and my recommendation is to look elsewhere if they don't offer you such a credential. Taking in advantage that, most of the scientific institutions in Arabic world can not offer payment for fresh graduates and they have limited funding, which made them in a big need for graduates who do not ask for payment.

7) Read even when it is difficult to understand; bioinformatics is a science that is changing every day and few textbooks will cover all of its aspects so you need to keep reading. That you will be able to understand a few more every time you start reading new research, finally you will get a clear understanding of most of these articles. While you're reading, take care of the software that they have used and try to run it using sample data. **Supplementary 1** contains some simple articles and reviews you can start with.

8) Self-learning is not the full answer; while self-learning through online resources such as YouTube (15) and Academia (16) is very helpful in expanding your knowledge, you need to deal with real problems by hand in order to sharpen your abilities, and this will not happen without joining a scientific group.

9) The way to learn more is to teach others; passing bioinformatics skills to others would open your eyes to different applications of the same tools and knowledge gaps, granting you more opportunities in the near future.

Operating systems and bioinformatics

Although most of the bioinformatics courses are interested in teaching software,I would suggest that it is not more about the tool than the environment in which you operate. Many scientists in our Arabic world use the Microsoft Windows Operating System every day to manage and evaluate their data. Microsoft's operating system is commercially closed and hackable (17). Such an environment is not a natural space for innovation and research, especially when dealing with mega-biological data.

The normal bioinformatics operating system should be opened (could be programmatically modified), highly secured, free of charge, and compatible with all bioinformatics software. Of which only one operating system could have granted these privileges, Linux. Linux is a Unix-like and often POSIX-compliant operating system (OS) based on the design and distribution of free and open source code. The basic element of Linux is the Linux kernel, the very first operating system kernel introduced by Linus Torvalds on September 17, 1991. The Free Software Foundation uses the name of GNU/Linux to describe a complex operating system (18).

Most Arab researchers fear Linux operating system, depending on the false belief that Linux is only a black and complicated command line window. On the contrary, Linux can operate videos, games and handle all the types of data files you use in windows, and it has free writing programs like Libreoffice (19) and a lot of wonderful programs and tools that could change your life. Linux also gives users a bit of control about what occurs on their machine what does and doesn't changing. You could learn more about the advantages of Linux over Windows through a number of articles (20).

Linux has different flavor distributions such as Ubuntu (21), Fedora (22), openSUSE (23), Red Hat (24) and many others. Mostly there is no difference between the various Linux distributions, although there is some variance between the software configuration , where it does not influence the core system performance or stability. Ubuntu is one of the most common Linux distributions (25), and the initial findings indicate that Ubuntu do's not require technical assistance (26),

has a defined and simple graphical user interface (GUI) and has been incorporated into some Arabic education systems (27). Ubuntu has a great and powerful version named Bio-Linux (28). In 2002, Dr. Dawn Field launched the Bio-Linux system under the NERC Environmental Bioinformatics Program (29). Bio-Linux release allows easy access to a versatile computing environment pre-loaded with bioinformatics software from basic data analysis tools to advanced analytic framework programming packages (21). You can download and install Bio-Linux in an ISO file format from the official website (http://environmentalomics.org/bio-linux/).

The only downside is that Bio-Linux distributions are a way behind Ubuntu's latest updates, and most hardware drivers such as WIFI, computer mouse, or touch pad, are not installed which involve certain installation skills. In order to overcome this issue, my recommendation is to install the latest Ubuntu release and then download Bio-Linux packages via the Linux Synaptic Package Manager. Alternatively, you can download the edition of Bio-Linux and then request the system to be upgraded with simple commands.

With the Ubuntu Software Center, you can quickly download and install thousands of Linux software without any complications. The Linux terminal is another way the software can be used (**Supplementary 2**). Although some tools require some command lines to be ready for analysis, most of them do not require installation skills. There are few Linux guidelines for fresh bioinformatics students:

1) Install and boot from Bio-Linux using flash memory for a short period of time while you are in practice (**Supplementary 2**).
2) Learn to use all the tools and software that Ubuntu has provided, even if they were really basic, such as LibreOffice and Calculator. This would eliminate the fear of using a new operating system to allow you more comfortable with the environment.
3) Bio-Linux has a web page documentation, and sample data for most its bioinformatics software, try to open and use it through the Desktop panel.
4) Try to write your own easy manual notes for all software built on Bio-Linux, explain the use, input, output and how you can use this tool in your future research. In **Supplementary 2**, I am explaining how to use a few of these tools.
5) After a few weeks try installing the latest version of Ubuntu or Bio-Linux , update and lunch Jemboss (30) using the Ubuntu Software Manager
6) Try to practice simple Linux commands and use as many tools as you can.

Computer science and bioinformatics

Most Arabic bioinformatics students prefer to start with programming languages such as Python for biological data analysis. Python has many advantages when it comes to bioinformatics and has several advantages as a computer language compared to languages such as PERL, which is also a famous bioinformatics computer language (31,32). The problem, though, is that most of these learners neglect and ignore the basic rule of computer language learning, which is the need to understand the basics of computing and its core structures, such as object orientation programming (OOP). On the other hand Java is a very common computer language in our Arabic country, and most commercial training companies offer paid courses to learn this language. My concern is that, although Java has a smaller relationship with bioinformatics compared to Python, few training companies offer python and fewer do so professionally in our Arab world compared to Java.

In this regard, while you are learning computer language, you need to understand the fundamentals of computer science. These fundamentals involve algorithms that are a sequence of commands typically used to solve a problem or perform a computation, and can be expressed in a finite number of steps and a well-defined formal language (33). The best way to learn algorithms in your first step is to learn algorithms by programming languages that you have selected to learn such as Java (34), Python (35), and PERL (36). Using this method, in addition to learning the algorithm effectively, you will be able to sharpen your coding skills.

After studying programming language in such a smarter way, you can easily add other programming languages to your set. Among these languages is R, which is a statistical programming language and common computational tool for data analysts, and is it has become one of the most commonly utilized programming languages in bioinformatics software. This is mainly due to its performance and the richness of the libraries that could be used for data manipulation and simulation (37). R is very simple, and there are large numbers of R libraries developed for bioinformatics. It might be difficult for biological researchers to learn coding and algorithms, and my recommendations for easy learning are:

1) Practice basic codes as much as you can, I would prefer to write these codes a couple times by hand and then run it under the computer language compiler.
2) Don't spend much of your time on graphical user interface (GUI) coding, most common bioinformatics software don't have a GUI, but instead focus on learning programming skills.
3) Do not forget that, for the purpose of interpreting biological data, you are studying programming language, and you are not a computer scientist. Based on this, don't go deep into programming language

learning, instead of considering what it can give you to achieve your goal.

4) After learning basics, you can search for programming codes that address simple biological issues such as DNA transcription and translation, reversing DNA sequence, and scripts which reflect the use of simple genetic concepts without the use of external libraries.

5) Do not use libraries/modules/packages for simple tasks in your first steps; you need to learn how to use computer language to apply scientific principles in algorithmic analysis.

6) Organize yourself; the first lines of your scripts should indicate the use of the script and its input and expected output data. You'll use these scripts in several occasions, and it's unfair to waste your time rewriting the same code.

7) You can use the same script or tool differently; this depends on your imagination.

8) Try to write clean and structured scripts where the overused programming functions can be found in well-organized libraries. It would provide the ability to use the same collection of libraries for various uses in different projects.

9) Backup your scripts every week.

Mathematics and bioinformatics

Most educational systems in our Arab world are concerned with teaching students without proper application of mathematics principles where student knowledge of these basics are not linked to their use. There is an easier way to teach mathematical rules in other countries by providing real-life experiments to learners (38). Most Arabic students have a decent mathematical background based on their high school education, which might be enough to start learning the science of bioinformatics, but my basic concern is that, they need to get missing knowledge through straightforward courses.

The main target is a sub-science of mathematics, statistics. Statistics is a systematic and random analysis of variability. In many aspects of scientific research, statistical methods are valuable. They constitute the research of the right way to collect process and interpret data (40). The connection between statistical methods and bioinformatics is very critical; you might argue that most of the biological data analysis could not be conducted without a good statistical background (41). As a bioinformatician, you will continue to learn the fundamentals and implementations of this science in biology for the most part of your life.

On the other hand, the basics of this science will need to be practiced in a simple way, where you can use R programming to study statistical methods. As I mentioned

earlier, R is a programming language that has been written for statistical analysis in general, and studying basic statistics through R could provide an opportunity to learn a new language and to understand the basics of statistics in the main time. This doesn't mean that you don't have to solve statistical problems by hand, but it does mean that you have to understand both ways.

Early publishing of students and graduates

Although there are concerns about early scientific publishing in the Arabic education system, it is very important for bioinformatics students to publish their research, scripts, and pipelines (series-connected processing systems in which the output of one component is received from another) during their early years. Some of these articles would not include much, but this is aimed at improving their interest in research. It would also give them the ability to sharpen their writing skills, address the comments of reviewers and the international research society. This would add to their computing skills in early age more specificity and professionalism. You should try to write your publications for the tools, scripts, pipelines or methods you have developed with or without other colleges and you need some guidelines on how to publish in this regard:

1) First of all, you need to know that most of the software articles are one or two pages long. It must include a brief introduction, a comprehensive methodology and some discussion.

2) Don't extend your article beyond two pages, the more you do, the more unstable your sentences will be.

3) Discuss the benefits of your tool, how you deal with the input, where adding a flowchart algorithm would be efficient.

4) Use simple English.

5) Cite previously published work by comparing your tool with others. It doesn't matter if your tool doesn't add much; the most important thing is that, it's your tool.

6) Start with small publications.

7) Search for free journals. In this regard, BioRxiv (42) is a non-profit, electronic archiving and distribution platform for the pre-printing of life science research papers. It was founded by the Cold Spring Harbor Laboratory, a scientific and academic organization.

Learn new programs and use published scripts

Hundreds of bioinformatics tools are released daily, specialized in the analysis of different data types, and written in a variety of programming languages. In order to use these tools, you need to take two steps: (1) test the configuration of the software using sample data and (2) read tutorials (if available) describing the different parameters and input data

that this tool could handle, which would enable you to understand and analyze the outputs of these tools. Using sample data and reading tutorials will save you time and provide you with a simpler way to resolve the error messages.

On the other hand script archives such as GithHub are the most important source for software scripts with more than 10 million repositories (43). Such websites offer you the ability to access open source software, script codes and pipelines in C, Python, PERL, R and other programming languages freely. You may be able to use such resources to know more about coding or bypass basic programming functions by using these codes as external libraries, but ethically you need to reference the published websites, articles or acknowledge the source of these scripts in your codes. It could be difficult to use these scripts and you need advice on how to handle it:

1) Begin with basic codes; you can find several tutorial codes for programming language implementations in bioinformatics, which you will find useful in sharpening your programming skills.
2) In order to understand any script, you need to divide the script parts according to what they do, use the print function to show the performance of the input processing stage after each step on the screen.
3) To understand any code structural algorithm, block, remove and replace any line debug (run by compiler) and monitor performance changes. Furthermore, the key to understanding the overall algorithm could be to comprehend how to evaluate script errors when ignoring certain coding lines.
4) Draw a basic flowchart that demonstrates your interpretation of how this program manages input data.

Start learning cloud computing

The management of large biological data requires sufficient computing power and storage capacity. Cloud computing can provide Arab research with large capabilities via open and often free websites. Some of these servers are Galaxy (44) and Cyverse (45). From my point of view, all these clouds and others provide computing resources and storage capacity to handle biological data, but I would prefer Cyverse more because it is more robust and offers a lot of software and simple GUI interfaces.

Using such clouds can offer you the ability to manage big information and to conquer the poor and slow speed of the Internet which inhibits you from accessing those data; you can do even more than that. Such as:

1) Cyverse gives you the ability to email your scripts to administration to handle or convert those scripts to a tool that you can use.
2) Using the tools available to build your own pipeline.

3) You can transfer or upload data from a private FTTP server such as NCBI (46), Ensemble (47) or from your computer.
4) Choose to use the storage capacity provided by these platforms to store data instead of overfilling your computer memory.
5) Such platforms have their own research communities, where you can submit your questions.

Life as a bioinformatician in Arabic countries

Many Arab countries do not support research (48) and fresh bioinformaticians may suffer from a lack of funding. My answer to this problem is to begin your career as a freelancer. Freelance employees are terms that are commonly used by a self-employed person who is not usually a long-term employer. Fresh bioinformatics graduates may offer their data analysis expertise to local and international research groups in exchange for money. In this manner, they will support their continuing self-education and their personal goals. Considering that, this path needs good research knowledge, free and low-cost services and communication skills. While freelance is a good way to start your life as a bioinformatician, some considerations are required:

1) Don't manipulate output reports for those who want to finish their dissertations without having worked.
2) Finish the work of each researcher without doing more than is necessary and not less than is essential.
3) Note that freelancing is not a type of research, and you need to write any scientific papers for yourself in order to build your own scientific background.
4) Once you have a successful and well-paid position, avoid freelancing and start sharing the credit for your work.

Bioinformatics and communication skills

Communication skills for bioinformatician are more than necessary and are very important for researchers, especially when working in scientific teams. Communicating with other bioinformaticians, biologists and researchers is critical to understanding, resolving and addressing daily bioinformatics issues. Such challenges could be software bugs, incorrect results and experimental methodologies. In most of our Arabic countries, there is a weak link between different scientific groups so, for order to overcome such difficulties, Arab bioinformaticians could use research communities to interact with each other and with international scientific teams. There are various advanced research, e.g. Ask Ubuntu (49), BIOINFORMATICS (50) and ResearchGate (51). Such groups will help young scientists start their careers and find new bioinformatics problems that they could address.

The future for you

Following basic learning, one or two programming languages and the handling of different biological data through a variety of analytical techniques, the expected next phases are as follows:

1) Learn C programming language; C is faster and consumes less memory than other programming languages. Although learning C is difficult, most famous bioinformatics programs such as the BLAST (52) software package are written in C. The problem is that Python, PERL and R are slower than C when handling large data, and then when performing complicated processes and memory management are needed (53). You will need to write some of your codes in C in the future to make your program faster and more efficient, even if these scripts need more lines to express. Like C, FORTRAN is a very important programming language, particularly when it comes to statistical analysis, and some may argue that it is faster than C (54).

2) Boost your understanding of SHELL scripting. SHELL is the programming language used by the Linux system (55), improving your reading and writing skills in such a language will enhance the management of complex tasks. In some cases, you can create hybrid scripts where you compose an algorithm to manipulate data throughout different computer programming languages such as Python, C, Julia and R, in addition allowing you the flexibility to use what the Linux system could provide.

3) Don't repeat your work; try to handle different types of data and use different methods of analysis to gain more knowledge and experience.

4) Machine Learning (ML) is a natural outgrowth of a combination of computer science and statistics and answers the question of how to build machines that automatically learn from experience (56,57). Understanding ML is very essential for your career and could aid your scientific background. Python, R, Java and C have specific libraries for ML, but Python has the maximum attention (58).

5) Begin the web programming learning process. Internet development is essential in order to make massive, definitive data available to the scientific community. In fact, you could create your own online software that would make it easier for scientists with poor programming skills to use. Python has its own web programming packages, such Django.

Mental and social life of Arabic bioinformatician

Research life is demanding, frustrating and highly competitive. Where postgraduates need high marks for their courses and a variety of publications under their belts if they want to earn excellent scholarships and research positions (59). It is important to take part in activities and social structures in order to promote the academic and social growth. Many current students in bioinformatics have been prepared to learn new research methodologies and programming languages and to ignore social life. Sadly, this kind of behavior is very risky with several articles on psychological science relating mental illness with creativity (60).

As a bioinformatician, you need to socialize with others and participate in outdoor activities to boost self-efficiency and promote team harmony. These type of activities are highly recommended in the field of mental health through a number of scientific articles (61). In fact, physical activity has been shown to be related to mental health and can play a key role in sustaining moderate to severe mental health conditions, especially anxiety (62). First of all, there is a high rate of depression and work stress, especially for those working as researchers in developing countries who need to maintain their mental health in order to thrive and thrive.

Conclusion

Bioinformatician is a person who has the ability to take advantage of three different sciences, to think innovative and to reshape complexity into simplicity. His capabilities rely on a sound scientific experience, a high level of knowledge in addition to patience, enthusiasm and productivity. Our Arabic world needs this kind of expertise in translating biological knowledge into practice, preserving our natural ecological resources and boosting our standard of living. As citizens, we can not change the education system, the economic climate or our society, but as researchers, we can inform others how to deal with these conditions and achieve their goals with less sacrifice as possible.

The secret of learning bioinformatics relies basically on the willingness of our students and researchers to acquire new sciences. Bioinformatics as a discipline is the most attractive and interesting field in which lack of resources, materials and equipment is not a concern, but requires additional knowledge and a high level of creativity that seems as easy to learn as it is difficult to obtain. My final advice is to be guided and to lead, both procedures could fill the gap in your scientific background, where opening your mind to new ideas would result in a better way, and deeper understanding and interaction with others would offer you multiple choices.

Acknowledgement

Supplementary Files

The Supplementary Material for this article can be found online at: **http://supplementary.highlightsin.org/20193_Sup.zip**

Supplementary file 1: Articles and reviews on bioinformatics could be used to initiate the learning process in this field.
Supplementary file 2: Short tutorial for some Bio-Linux basics.

References

1. Welch L, Lewitter F, Schwartz R, Brooksbank C, Radivojac P, Gaeta B, et al. Bioinformatics curriculum guidelines: toward a definition of core competencies. PLOS Comput Biol. 2014;10(3):e1003496.

2. Bansal AK. Bioinformatics in microbial biotechnology-- a mini review. Microb Cell Fact. 2005;4(1):19.

3. Escobar-Zepeda A, de León A, Sanchez-Flores A. The road to metagenomics: from microbiology to DNA sequencing technologies and bioinformatics. Front Genet. 2015;6:348.

4. Larson RS, Oprea TI. Bioinformatics and Drug Discovery. Springer; 2006.

5. Fernald GH, Capriotti E, Daneshjou R, Karczewski KJ, Altman RB. Bioinformatics challenges for personalized medicine. Bioinformatics. 2011;27(13):1741–8.

6. Larranaga P, Calvo B, Santana R, Bielza C, Galdiano J, Inza I, et al. Machine learning in bioinformatics. Brief Bioinform. 2006;7(1):86–112.

7. Omondiagbe DA, Veeramani S, Sidhu AS. Machine Learning Classification Techniques for Breast Cancer Diagnosis. In: IOP Conference Series: Materials Science and Engineering. 2019. p. 12033.

8. Villanueva-Meyer JE, Chang P, Lupo JM, Hess CP, Flanders AE, Kohli M. Machine learning in neurooncology imaging: from study request to diagnosis and treatment. Am J Roentgenol. 2019;212(1):52–6.

9. Ballester PJ. Machine Learning for Molecular Modelling in Drug Design. Multidisciplinary Digital Publishing Institute; 2019.

10. Yamamoto K. Distillation of crop models to learn plant physiology theories using machine learning. PLoS One. 2019;14(5):e0217075.

11. Filippi P, Jones EJ, Wimalathunge NS, Somarathna PDSN, Pozza LE, Ugbaje SU, et al. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precis Agric. 2019;1–15.

12. Zhi Q, Saitoh T, Nakajima M, Saitoh T. A development of content-based video summarization system using machine-learning and its application to analysis of livestock behavior. In: International Workshop on Advanced Image Technology (IWAIT) 2019. 2019. p. 1104911.

13. Hadidi MS, Al Khateeb JM. Special education in Arab countries: Current challenges. Int J Disabil Dev Educ. 2015;62(5):518–30.

14. Loucif S. Bioinformatics in Middle East program curricula—a focus on the Arabian Gulf. Educ Inf Technol. 2014;19(2):375–82.

15. YouTube LLC. YouTube. Retrieved. 2011;27:2011.

16. Academia. Academia. Available from: https://www.academia.edu/

17. Dragoni N, Giaretta A, Mazzara M. The internet of hackable things. In: International Conference in Software Engineering for Defence Applications. 2016. p. 129–40.

18. Zastrow R. Linux Essentials. 2016;

19. Gamalielsson J, Lundell B. Sustainability of Open Source software communities beyond a fork: How and why has the LibreOffice project evolved? J Syst Softw. 2014;89:128–45.

20. Browder T. Why Linux? 2018;

21. Ubuntu. No Title. Available from: https://ubuntu.com/

22. Fedora. No Title. Available from: https://getfedora.org/

23. openSUSE. No Title. Available from: https://www.opensuse.org/

24. Red Hat. No Title. Available from: https://www.redhat.com/en

25. Saeki T, Nishiwaki Y, Shinagawa T, Honiden S. Bash on Ubuntu on macOS. In: Proceedings of the 8th Asia-Pacific Workshop on Systems. 2017. p. 17.

26. Dawson M, DeWalt B, Cleveland S. The case for UBUNTU Linux operating system performance and usability for use in higher education in a virtualized environment. 2016;

27. EL-Tantawy SA. Ubuntu. Port Said University; 2016.

28. BioLinux. No Title. Available from: http://environmentalomics.org/bio-linux/

29. NEBC. No Title. Available from: http://nebc.nerc.ac.uk

30. Carver T, Bleasby A. The design of Jemboss: a graphical user interface to EMBOSS. Bioinformatics. 2003;19(14):1837–43.

31. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25(11):1422–3.

32. Langtangen HP, Barth TJ, Griebel M. Python scripting for computational science. Vol. 3. Springer; 2006.

33. Moschovakis YN. What is an algorithm? In: Mathematics unlimited—2001 and beyond. Springer; 2001. p. 919–36.

34. Lafore R. Data structures and algorithms in Java. Sams publishing; 2017.

35. Chou PH. Algorithm education in Python. Proc Python. 2002;10:2.

36. Orwant J, Hietaniemi J, Macdonald J, MacDonald JD. Mastering algorithms with Perl. " O'Reilly Media, Inc."; 1999.

37. Gentleman R. R programming for bioinformatics. Chapman and Hall/CRC; 2008.

38. Okigbo EC, Osuafor AM. Effect of using mathematics laboratory in teaching mathematics on the achievement of mathematics students. Educ Res Rev. 2008;3(8):257–61.

39. He M, Petoukhov S. Mathematics of bioinformatics: Theory, methods and applications. Vol. 19. John Wiley & Sons; 2011.

40. Chatfield C. Statistics for technology: a course in applied statistics. Routledge; 2018.

41. Jiang R, Zhang X, Zhang MQ. Basics of Bioinformatics. Springer; 2016.

42. bioRxiv. bioRxiv. Available from: https://www.biorxiv.org/

43. Kalliamvakou E, Gousios G, Blincoe K, Singer L, German DM, Damian D. The promises and perils of mining GitHub. In: Proceedings of the 11th working conference on mining software repositories. 2014. p. 92–101.

44. Afgan E, Baker D, Coraor N, Goto H, Paul IM, Makova KD, et al. Harnessing cloud computing with Galaxy Cloud. Nat Biotechnol. 2011;29(11):972.

45. CyVerse. CyVerse. Available from: https://cyverse.org/

46. Maglott D, Ostell J, Pruitt KD, Tatusova T. Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res. 2010;39(suppl_1):D52--D57.

47. Hubbard T, Barker D, Birney E, Cameron G, Chen Y, Clark L, et al. The Ensembl genome database project. Nucleic Acids Res. 2002;30(1):38–41.

48. Abu-Orabi ST, General S. Scientific research & higher education in the Arab world. In: CLICKS MENA Higher Education Leadership Forum. 2013. p. 12–3.

49. Ask Ubunut. Ask Ubunut. Available from: https://askubuntu.com/

50. Bioinformatics. No Title. Available from: https://www.bioinformatics.org/

51. Yu M-C, Wu Y-CJ, Alhalabi W, Kao H-Y, Wu W-H. ResearchGate: An effective altmetric indicator for active researchers? Comput Human Behav. 2016;55:1001–6.

52. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389–402.

53. Prechelt L. An empirical comparison of seven programming languages. Computer (Long Beach Calif). 2000;33(10):23–9.

54. Robison AD, Dubois PF. C++ gets faster for scientific computing. Computers in Physics. 1996 Sep;10(5):458-62.

55. Arthur LJ. UNIX shell programming. John Wiley & Sons, Inc.; 1990.

56. Mitchell TM. The discipline of machine learning. Vol. 9. Carnegie Mellon University, School of Computer Science, Machine Learning; 2006.

57. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science (80-). 2015;349(6245):255–60.

58. Müller AC, Guido S, others. Introduction to machine learning with Python: a guide for data scientists. " O'Reilly Media, Inc.; 2016.

59. Ortega A. Three tips to achieve academic success — while enjoying a social life. Nature. 2019

60. Nettle D. Schizotypy and mental health amongst poets, visual artists, and mathematicians. J Res Pers. 2006;40(6):876–90.

61. Fernee CR, Gabrielsen LE, Andersen AJW, Mesel T. Therapy in the open air: Introducing wilderness therapy to adolescent mental health services in Scandinavia. Scand Psychol. 2015;2.

62. Paluska SA, Schwenk TL. Physical activity and mental health. Sport Med. 2000;29(3):167–80.