A Machine Learning Approach for Prediction of Students' Admissibility for Post-Secondary Education using Artificial Neural Network

Anietie Ekong
Department of Computer Science,
Akwa Ibom State
University, Nigeria

Abasiama Silas
Department of Computer Science,
Akwa Ibom State
University, Nigeria

Saviour Inyang
Department of Computer Science,
Akwa Ibom State
University, Nigeria

ABSTRACT

Student admission's process is a method of selecting qualified candidates for admission. Challenges such as facility constraints and insufficient ability to meet the continuously rising needs of post-secondary education. There is still an absorption capacity problem in some parts of the world as the growing number of students applying for admission for post-secondary education far surpasses the rate of expansion and this makes the selection process to be a daunting tasks. In this study, Artificial Neural network (ANN) was adopted for the determination of admissibility of candidates for post-secondary education based on (O'level Results, CGPA (Cumulative Grade Point Average), Departmental Rank (DPR) etc. Results indicated effective prediction based the performance analysis using the Confusion Matrix and AUC - ROC and gave a 99% accuracy on the dataset.

Keywords
Machine Learning, Neural Network, Model, Prediction, Student's Admission.

1. INTRODUCTION

Student admissions process is a method of choosing students for higher education [1]. The current system for students' admission at various universities consists of a series of processes beginning with registration, determination of a student cutoff point in post-secondary exams, grades of students from secondary level education, the strata from which the student hails and the capacity of the school with respect to resources, both human and non-human. The major educational challenge of today is how to further educate more people, with the majority of them completing postsecondary education [2]. University education is no longer merely the quickest path to a middle-class lifestyle; it seems to be and is perceived by many to be the only option [3]. Challenges such as financial, facility constraints, low standards, and insufficient ability to meet the expanding demand for postsecondary education have accompanied this continuously rising need for post-secondary education [4,5]. Considering the significance of education and its role in the development of the society, ensuring that qualified candidates are admitted using a process that is fair is necessary.

Artificial Neural Network being a subset of machine learning and at the heart of deep learning algorithms is an efficient approach in handling a lot of problems including ability to efficiently predict a student's admissibility based on learned data. ANNs are inspired by the human brain, mimicking the way that biological neurons send signals to one another. This research leverages on this machine learning algorithm for the prediction of admissibility of candidates that are considered for post-secondary education based on relevant metrics.

2. LITERATURE REVIEW

2.1 Overview of Higher Education

Today, education is a crucial resource for any nation's economic, social, and intellectual growth and the role of university education in supporting these indices cannot be overstated [6]. In order to select which individuals will be admitted, universities look at a variety of factors in the prospective students. Nigerian university education has been unable to match the demand for increased university admissions and secondary school graduates and other members of society who are undereducated have an overwhelming need for higher education beyond what the government can give [7]. It is a well-known fact that Nigeria's population has been significantly increasing over time. The nation population was predicted to be 200,963,599 in 2019, with a 2.60 percent growth rate. [8]. According to the World Education News and Reviews, Western Education Service [9], over 60% of Nigerians are under the age of 24. This demonstrates the rapid growth of the youth population. To meet the educational needs of this rapidly growing student population, the government has always worked to increase the number of universities, polytechnics, monotechnic, and colleges of education[9]. However, despite the federal and state governments' expansion strategy, admission into post-secondary schools has become extremely competitive, as the increased demand has remained relentless over the years.

2.2 Machine Learning in Student Admission

Higher education has become a necessary and fundamental aspect of most individual's preparation for a career. However, for university education to fulfill its mission, students must be placed in the appropriate institution and /or discipline [10]. In the digital age, college and universities admission have evolved tremendously. Students and staff benefit from the integration of innovation and learning. Artificial Intelligence (AI), Cloud computing, mobile computing, and digital technologies are hastening the transformation of education system around the world. Artificial intelligence (AI) is a key topic in science and a technological in the 21st century. It uses technology to create an intelligent idea that is akin to that of the human brain. Various industries have benefited from incorporating AI into their day-to-day business operations and educational sector cannot not be excluded. Machine Learning (ML) is a branch of artificial intelligence that detects complicated and unseen patterns or knowledge in massive volumes of data and then makes intelligent judgments based

on that data [11]. One of the most important components of machine learning is a training model that uses a variety of dependent and independent variables and is reliant on the types of learning algorithms used (supervised, semi supervised, or unsupervised). Machine learning (ML) is often used to make predictions in a variety of fields, such as [12], credit default prediction in bank [13], cyber security [14] etc.

In recent years, there has been a surge in research interests in the use of Machine Learning in education, especially within institutions of higher learning. According to a recent survey, educational decisions are generally made primarily on the perceptions and experiences of school management stakeholders/students, instead of knowledge-rich data. By making relevant instructional material readily available to students and other individuals, Machine learning (ML) approaches have also been proved to be effective in enhancing outcomes at numerous educational establishments [15,16]. With regards to predicting and detecting quality educational related problems for students and judgment call, as well as improving other management solutions linked to optimizing students' demands, machine learning is becoming more widespread and important in educational contexts. Numerous studies have also been conducted on Machine Learning (ML) approaches that have been used to predict admission to universities,student allocation and admittance into their

selected universities which are presented in [17, 18, 19]. Furthermore, according to [20], student admittance has been predicted with a high degree of accuracy using machine learning and predictive modeling. The strategy isn't specific to the institution in this case. Data submitted by the applicant in the containing a large number of data necessitates further investigation in order to predict student'sadmissibility. Also, in [10], different understandable machine learning approaches such as Decision Tree, Random forest, Extra tree classifiers [ETC], Gradient boosting classifiers, and Support Vector Machine (SVM) were evaluated in their study to predict students' likely field of specialization before enrollment at the undergraduate level based on existing employment market and knowledge.

3. RESEARCH METHODOLOGY

In this study a Neural Network Machine Learning model is adopted in the prediction of student's admission in a datasets. Artificial Neural Networks (ANNs), sometimes known as Neural Networks (NNs), are computer systems designed after the biological neural networks seen in mammalian brains. Artificial neurons are a network of interrelated units or nodes that resemble neurons in a biological brain. Each link, like synapses in the human brain, can transmit a signal to other neurons. An artificial neuron receives a signal, analyzes it, and then transmits it to the neurons to which it is attached. The output of each neuron is calculated using a non-linear function as its inputs, and the "signal" at a connection is a real number [20]. Nevertheless, there are numerous types of neural network, ranging from perceptron, multilayered perceptron, feedforward, recurrent, and convolution etc. this research will adopt a feedforward neural network which entails an algorithm that is biologically inspired. It is made up of a (potentially large) number of basic neuron-like processing units that are layered together. Every unit in a layer is linked to all the units in the layer before it. These connections are not all created equal: each one may differ in terms of strength or weight.

The weights on these connections represent a network's knowledge. The units of a neural network are frequently referred to as nodes. Also, a sigmoid activation function will be adopted to "fire" the neurons to the output, this is depicted in equation 1.

$$f(x) = \frac{1}{1 + e^{-x}} \qquad 1$$

This Sigmoid function accepts any real value as input and returns a value between 0 and 1. As illustrated figure1, the greater the input (more positive), the closer the output to 1.0, and the smaller the input (more negative), the closer the output to 0.0.
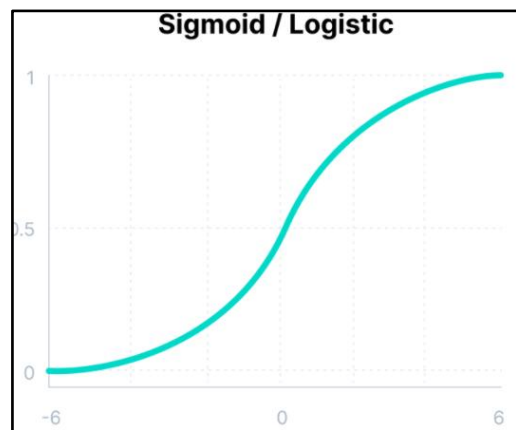

Figure1: Sigmoid activation function

3.1 Description of Key Components

The key system components are as illustrated in figure 2 and the details are as follows;

1. **Data collection:** Data collection is the systematic process of acquiring and measuring information on variables of interest in order to answer research questions, test hypotheses, and evaluate outcomes, in this research data was collected from online sources.

2. **Data Preprocessing:** The process of converting raw data into a comprehensible format is known as data preparation. This is done before using machine learning or data mining methods to ensure that the data to be used are of good quality.

3. **Model training and testing**: model training is used to fit the machine learning model, testing is to estimate the machine learning model's performance on new data that was not used to train the model.

4. **Evaluation and performance measure:** this section will carry out evaluation on the result of the prediction with neural network and the performance evaluation of the model using confusion matrix.
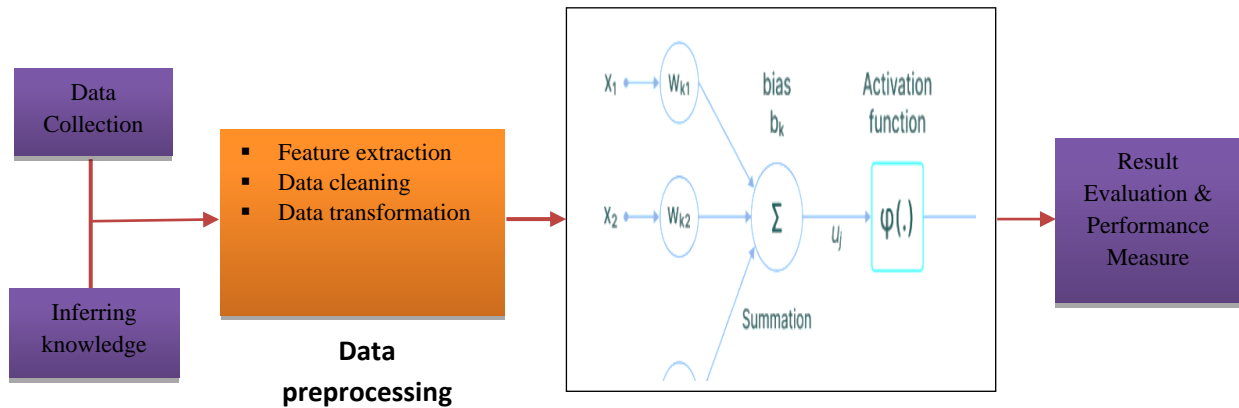
Figure2: Depiction of the phases in the proposed system

3.2 Problem definition

Modelling post-secondary admission is a very important measure when considering the rate of students or individuals that are seeking admission into these institutions across the globe. There are different processes that are involved in the admission processes of prospective student applying to university and there are basic requirement that each individual or candidate must meet. Hence, to model the admission of an individual candidate, certain criteria were taken into considerations. In this research, metrics such has Cumulative Grade Point Average, the rank of the department the student isapplying for, the results obtained from the secondary school that the student attended were all considered in the dataset which will be used for the prediction in order to ascertain whether the student is admissible not. These will serve as input variables into the neural network model for the purpose of training and testing and prediction.

3.3 Model Design

The Following are the steps adopted to design the Neural Network model of the proposed System.

    i.   Pre-process the data set.

    ii.   Check for outliners and inconsistencies in the dataset

    iii.   Normalize of the data

    iv.   Split the data set into training and test set in the ration of 70: 30%

    v.   apply a neural model in the trained set

    vi.   apply the test set on the trained model

    vii.   carry out performance evaluation on the model using confusion matrix

Furthermore, the dataset is depicted in the figure 3.

	admit	gre	gpa	dp	Olevel
1	0	380	3.61	3	0
2	1	660	3.67	3	1
3	1	800	4.00	1	1
4	1	640	3.19	4	1
5	0	520	2.93	4	0
6	1	760	3.00	2	1
7	1	560	2.98	1	1
8	0	400	3.08	2	0
9	1	540	3.39	3	1
10	0	700	3.92	2	0
11	0	800	4.00	4	0
12	0	440	3.22	1	0
13	1	760	4.00	1	1
14	0	700	3.08	2	0
15	1	700	4.00	1	1
16	0	480	3.44	3	0

Showing 1 to 16 of 400 entries, 5 total columns

Fig 3: Cross-section of the data set

4. RESULT & PERFORMANCE EVALUATION

This study uses a neural network method to predict whether or not a particular candidate requesting for admission will be accepted or not. Figure 4a,b,c. represents input before normalization.
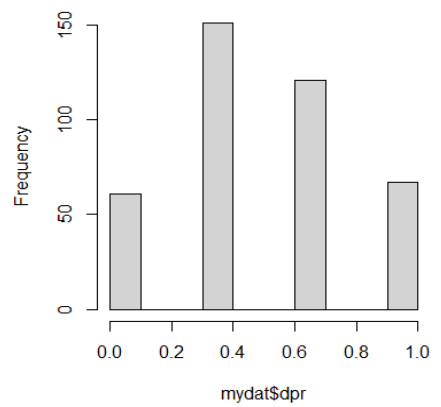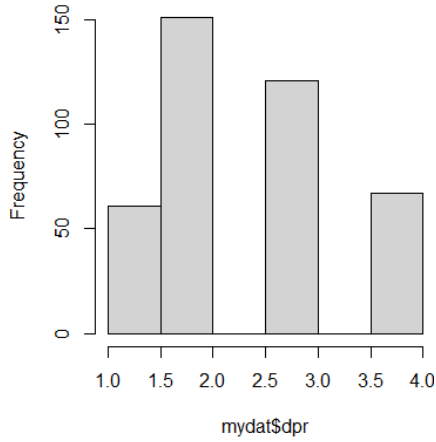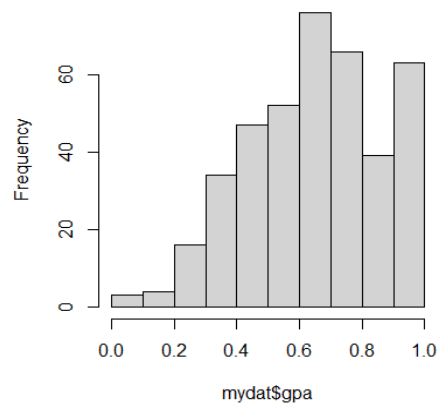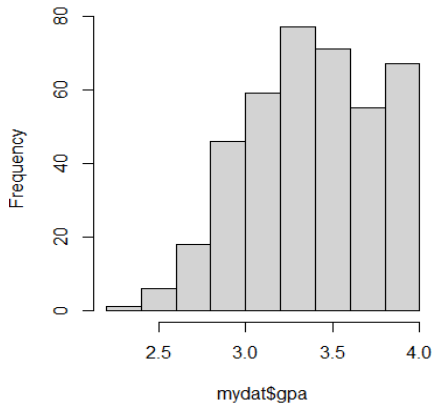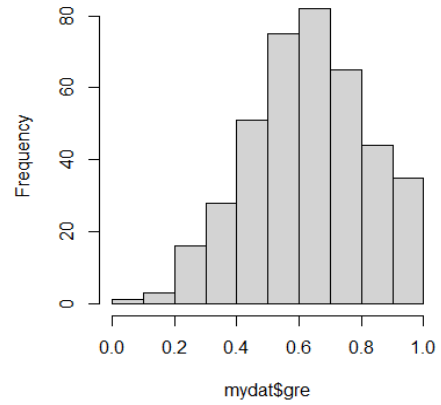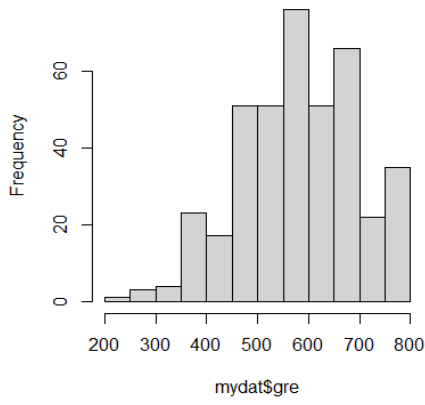
Figure 4a,b,c:Input before normalization

Furthermore, figure 5 a,b and c depicts input after normalization

Figure 5 a,b, c: Input after normalization

The architecture of the neural network training model is depicted in figure 6, this shows how the neural network model fits into the data set.
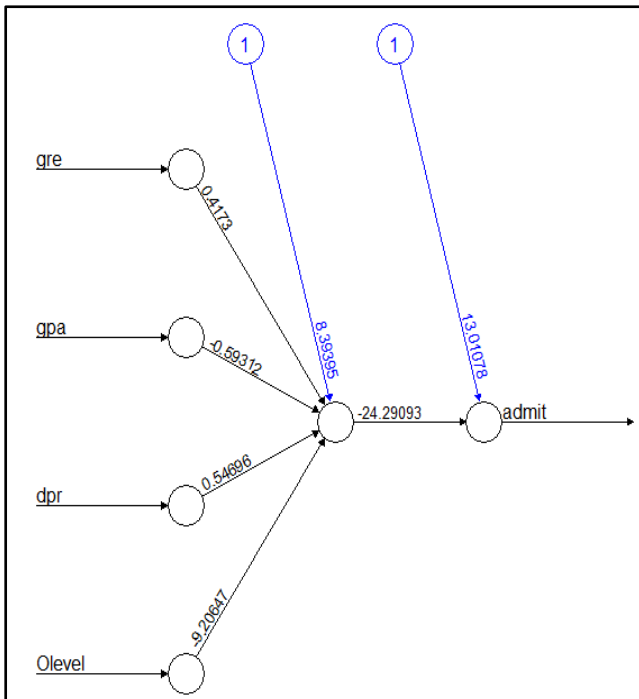
**Figure 6: Architecture of the neural network training model**

Furthermore, the confusion matrix is as presented in table 1;

Table 1: Confusion matrix

Predicted	Actual value	
value	0	1
0	188	0
1	1	92

The confusion matrix shows that out of a total of 281 observations in the training set, 188 were correctly predicted as student that were not admissible and 1 student was mis-predicted as admissible and 92 were correctly predicted as admissible and there were no mis-prediction. Based on the confusion matrix, the accuracy measure is calculated by finding the sum of the diagonals in the confusion matrix and divide it by the total observation in the matrix i.e. (sum (diagonals (confusion matrix) / sum (confusion matrix)) which the result gave a 0.9964413~99% accuracy.

Furthermore, another important method or metrics performance evaluation or analysis of model is receiver operating characteristic (ROC) graph or curve, which is a graphical representation of a binary classifier system's diagnostic performance as its discrimination threshold is changed. In this research, ROC graph and area under the curve (AUC) also enables proper evaluation of the model performance as shown in figure 7.
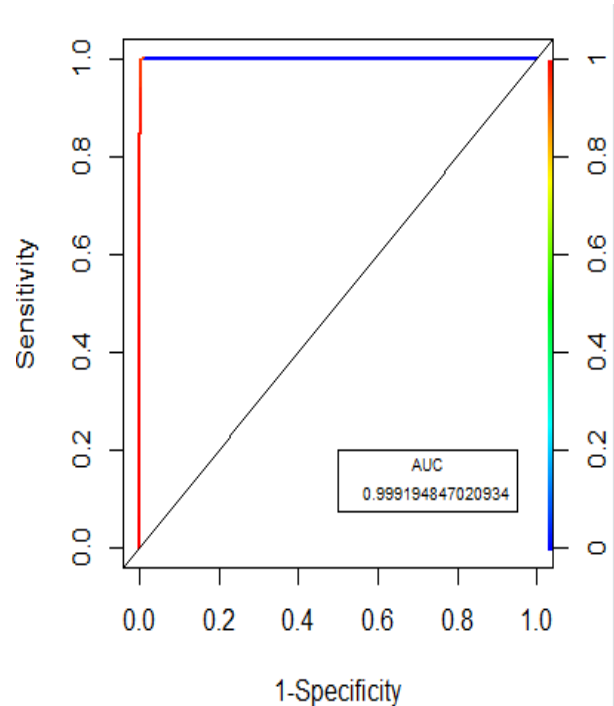
5. CONCLUSION

In today's world, the academic environment today cannot be left out in the use of efficient machine learning algorithms to aid in important management decisions such as student's admissions and admissibility. One of today's key educational challenges is to select the most suitable individuals to fill the limited spaces in post-secondary institutions. Devising the right and an efficient approach for admitting the enormous number of candidates that apply for admission into different institutions of higher learning is a key challenge. To this end, this research has established a machine learning framework to predict student admissibility into these institutions based on the metrics which was identified in the dataset. In this study, Artificial Neural network was adopted for the determination of admissibility of candidates for post-secondary education based on (Secondary(High) School Results, CGPA (Cumulative Grade Point Average), Departmental Rank (DPR) etc. Results indicated effective prediction based on the performance analysis using the Confusion Matrix and AUC - ROC and gave a 99% accuracy on the dataset

6. REFERENCES

[1] Hanifa F and H Ajie, (2019), "Design and Development of Student Admission System in the State University of Jakarta" in 3rd UNJ International Conference on Technical and Vocational Education and Training 2018, KnESocial Science, pp. 92–101.

[2] Hunt J. B. (2006). Educational leadership for the 21st century in American higher education: How does it measure up for the 21st century? San Jose, CA: National Center for Public Policy and Higher Education.

[3] Tierney T. J. (2006, May). How is American higher education measuring up? An outsider's perspective. In American Higher Education: How does it Measure up for the 21st Century? San Jose, CA: The National Center for Public Policy and Higher Education.

[4] Okebukola, P. (2010). Fifty years of higher education in Nigeria: Trends in quality assurance. Paper presented at

the 50th Independence Anniversary Conference, New York.

[5] Sawyerr, A. (2004). Challenges facing African Universities: Selected issues, African Studies Review, Vol47, No. 1, pp. 1-59.

[6] Ajayi, I. A., &Ekundayo, H. T. (2008, December). The deregulation of university education in Nigeria: Implications for quality assurance. Nebula, Vol. 5, No. 4.

[7] VinellaOkonta (2020). Trend Analysis of the Demand and Supply of University Admission Places in Delta State University (2013 – 2017). Universal Journal of Educational Research Vol8, No 12B, pp. 8118-8128, 2020.

[8] Worldometer, 2020.www.worldometers.info.

[9] BolajiF(2020). List of all Universities in Nigeria approved by NUC – 2020 Latest List. https://campusbiz.com.ng>list-ofuniversities. Retrieved on 17th June, 2020.

[10] Alsayed, A.O., Rahim, M.S..; AlBidewi, I.; Hussain, M., Jabeen, S.H., Alromema, N., Hussain, S.,Jibril, M.L.(2021) Selection of the Right Undergraduate Major by Students Using Supervised Learning Techniques. Appl. Sci. 2(11), 10639.

[11] Sharma, S.; Gupta, Y.K. (2021) Predictive analysis and survey of COVID-19 using machine learning and big data. J. Interdiscip. Math. 24, 175–195.

[12] Gogas, P.; Papadimitriou, T.; Agrapetidou, A(2018). Forecasting bank failures and stress testing: A machine learning approach. Int. J. Forecast. Vol34, pp 440–455.

[13] Alam, T.M.; Shaukat, K.; Hameed, I.A.; Luo, S.; Sarwar, M.U.; Shabbir, S.; Li, J.; Khushi, M. An investigation of credit card default prediction in the imbalanced datasets. IEEE Access 2020, 8, 201173–201198.

[14] Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Xu, M (2020). A survey on machine learning techniques for cyber security in the last decade. IEEE Access 2020, 8, 222310–222354.

[15] Dhar, J.; Jodder, A.K(2020). An Effective Recommendation System to Forecast the Best Educational Program Using Machine Learning Classification Algorithms. Ingénierie Des Syst. EgraveMes Inf. Vol 25, pp 559–568.

[16] Meng, Y.; Fu, M (2020). CMRS: Towards Intelligent Recommendation for Choosing College Majors. In Proceedings of the 2020 4th International Conference on Advances in Image Processing, Chengdu, China, pp. 152–157.

[17]Tan, L.; Main, J.B.; Darolia, R. Using random forest analysis to identify student demographic and high school-level factors that predict college engineering major choice. J. Eng. Educ. 2021, Vol. 110, pp 572–593.

[18] Awaliyah, M.M.; Kurniawati, A.; Rizana, A.F(2020). Profile matching for students specialization in industrial engineering major. IOP Conf. Ser. Mater. Sci. Eng., 830, 032063.

[19] Pertiwi, D.A.; Daniawan, B.; Gunawan, Y(2019). Analysis and Design of Decision Support System in Major Assignment at Buddhi High School Using AHP and SAW Methods. Tech-E, Vol 3, pp 13–21.

[20] KanadpriyaBasu, TreenaBasu, Ron Buckmire, and NishuLal(2019), "Predictive Models of Student College Commitment Decisions Using Machine Learning," Vol. 4, issue 2, pp. 1-18.

[21] Zou J, Han Y, So SS (2008). Overview of artificial neural networks. MethodsMol Biol. 2008;458:15-23. doi:10.1007/978-1-60327-101-1_2. PMID: 19065803.