

Comparison of Dense Net and over Logistic Regression in Predicting Leukemia Classification with Improved Accuracy

Barinepalli Vamsi kumar^{1,*}, S.Stella Jenifer Isbella²

¹Research Scholar, ²Project guide, Department of Medical Instrumentation, Saveetha School of Engineering, saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India,

Abstract: This study compares the performance of densenet and support vector machines (SVMs) in the diagnosis of leukemia disease, with the aim of improving the accuracy of the classification results. **Materials and Method** The Kaggle website is where the dataset was found. The dataset consists of 20 samples per group in JPG files with a resolution of 96 dpi and 512×512 pixel size. The sample size is determined using a pretest power of 80%, a threshold of 0.05, and a confidence interval of 95%. **Results:** For leukemia, dense net is 96.5%, whereas logistic regression is 89%. The significance levels for Densenet and logistic regression are data with $p=0.000$ ($p<0.05$) statistical significance difference respectively. **Conclusion:** Based on the findings, I believe that densenet performs superior to logistic regression.

Keywords: Novel-Accuracy, Dense Net, Disease, Leukemia, Logistic Regression, Machine Learning, Red Blood Cells, White Blood Cells.

INTRODUCTION

The blood cancer affects B lymphocytes, which are white blood cells and play a significant role in the immune system. leukemia chronic lymphocytic (CLL). The majority of CLL patients are older, with a median diagnosis age of about 70 (Baliakas et al. 2019). In recent years, there has been a dramatic improvement in the specificity and sensitivity of the detection of malignant cells in patients who by traditional measures appeared to be cured or in remission. This is possible thanks to advancements in technology and the rapid growth of our knowledge of the etiology of hematological malignancies (Brauer, Castillo-Chavez, and Feng 2019). How well a patient is responding to treatment can now be determined by tracking minimal residual disease (MRD), which has a sensitivity of 110⁻⁴ and may detect

* Corresponding author: bitmist2017@gmail.com

cancer cells at preclinical levels. Leukemia is brought on by bone marrow-derived blood stem cells. A soft, spongy substance called bone marrow makes up the majority of bones. The bone marrow produces blood cells. Myeloid stem cells and lymphoid stem cells are the two types of blood stem cells that develop. Red blood cells, which assist deliver oxygen, white blood cells, which help fight infection, and platelets, which aid in blood clotting, make up the majority of the cells in blood. The bone marrow creates hundreds of billions of fresh blood cells each day. Red blood cells are often generated and present in higher numbers. This essay explores how the bone marrow creates blood cells using mathematical methods (hematopoiesis) (Ahern 2018). The model consists of two age-structured partial differential equations. We time-integrate these equations to obtain a system of two nonlinear differential equations with scattered time delays, which represents the length of the cell cycle. The development of every cell historically This concept explains opulations. By developing a Lyapunov functional, it is demonstrated that the trivial equilibrium, if it is the only equilibrium, is globally asymptotically stable (Dillig and Tasiran 2019). A Hopf bifurcation is used to demonstrate that the nontrivial equilibrium, which has the most biological significance in machine learning, is unstable. Numerical simulations are used to illustrate the analytical conclusions (Rodríguez 2021).

55 publications in Google Scholar, 40 on the Web of Science. This essay will review current theories about the origins of acute myeloid leukemia in individuals with autoimmune disorders. We'll look at probable leukemogenic pathways in these patients as well as information indicating a relationship between their level of primary immunosuppression and the medications they used white blood cells. To make the greatest use of the available therapeutic choices, patients with chronic myeloid leukemia disease must integrate clinical and laboratory monitoring. Analyzing the molecular reaction. Real-time quantitative polymerase chain reaction is the most precise method for determining the effectiveness of tyrosine kinase inhibitor (TKI) medications (MR). Along with "major molecular response," which has emerged as a safe haven for survival since the initial trials of first-line imatinib treatment, the phrases "early molecular response" and "deep molecular response" have also recently been developed. It is thought that reaching these MR objectives during therapy within predetermined time intervals will result (Soverini et al. 2016). In a study of 1 1 5 individuals, including cases of lympho-sarcoma cell leukemia illness, Gray et al. showed that cases presenting with diffuse bone marrow infiltration had a poor result in comparison to cases presenting with a nodular or mixed (nodular and diffuse) pattern. In a preliminary investigation, several of us identified four different forms of CLL involvement in the bone marrow the white blood cells interstitial, nodular, mixed (interstitial and nodular), and diffuse. 2 There was a significant connection between these patterns and clinical phases, according to machine learning (Rozman et al. 1984).

Severe cardiovascular disease was a requirement for exclusion, although prior MDS treatment was not (except standard chemotherapy or treatment with a hypomethylating agent). Each trial site's regional ethics committee approved the study, which was conducted with all patients' written informed permission and in compliance with the Declaration of Helsinki (Lübbert et al. 2011). When the increased B-lymphocyte samples were tested by flow cytometry using CD5-FITC and CD20-phycoerythrin (PE) antibodies, the purity was 93%. (data not shown). Only living cells can be bound by DotScan microarrays. The inclusion of the microbeads had no effect on the capacity of the pure cells to be captured on microarrays ("[No Title]" n.d.). shows typical primary implantation locations that were employed in the current investigation. Photomicrographs of samples collected from rhesus monkeys are used to illustrate the immunolocalization of LIF, IL-1a, IL-1b, and IL-6 in cells of the conceptus and endometrium, respectively, at the primary implantation site (Dhara 2001). The conceptus and maternal endometrium at the site of implantation during the lacunar and villous stages, respectively, were scored using immunohistochemistry staining in a variety of cell types. It is critical to evaluate the degree to which cancer

survivors experience both positive and negative quality-of-life outcomes, as well as the psychosocial factors that are associated with these outcomes, in the context of treating and supporting cancer patients and survivors along a continuum of care that integrates psychosocial and medical support. Our increased understanding of quality of life and the elements that affect it should enable future research into which psychological support interventions should be given to which patients when during the first phases of diagnosis or therapy or over the long haul. That is crucial.

MATERIALS AND METHODS

The test was administered in the academic simulation lab at Chennai's Saveetha School of Engineering, Saveetha Institute of Medicine and Technical Sciences. The sample size for this article was estimated using clinical.com based on the findings of previous research, with an alpha error-threshold of 0.05, an enrollment ratio of 0:1, a confidence interval of 90%, and a power of 80%. Group 1 displayed a dense net (N=20), whereas group 2 displayed logistic regression (N=20). For this experiment, a total of 40 samples were used in machine learning.(Chabner and Longo 2018)

Data is taken from the kaggle website(Sehgal 2017). Samples are runned under Python 3.10 Software with Densnet and Logistic Regression(Lantz 2013). Using algorithms based on various cell characteristics, the cells are separated from the background (such as shape, color, and inner intensity). The taxonomy of cells separates red and white cells. The classifier might use color data to look for the nucleus. You can identify lymphocytes from other types of white blood cells by looking at the nucleus (e.g., a highly stained nucleus that may be located eccentrically, or a little amount of cytoplasm). Detection of potential lymphoblasts: In a group of lymphocytes, one can spot potential lymphoblasts by observing the morphological distortions of the cells.Leukemia microenvironment is a factor in investigations on medication screening.

The binary classification of all and HEALTHY classes using augmented data is where we performed the best white blood cells. The classification of the four leukemia subtypes into many categories received the second-best score. However when multiple classifications are adhered to and participated in the classification process, binary classification surpasses multi-classification because it enables the model to distinguish between the classes more clearly. In terms of Novel- accuracy and loss measurements, we found that SGD optimizer performs noticeably better than ADAM optimizer. More epoch iterations do not improve the model's performance because it outperforms rug screening tests in terms of outcome.

Our proof that the basic life characteristics of the stromal cells are unchanged by the hypoxic co-culture conditions lends credence to the proposed paradigm's validity. Also, we observed how the stroma in this experimental system prevents imatinib-induced cell death from killing leukemia cells, which leads to the development of treatment resistance and the progression of the disease.studies on screening rugs.

Overall, we suggest using this flow cytometric method to quickly evaluate the health of stromal cells in co-culture hypoxia models. The suggested strategy may usually be used to advance investigations into the relationship between leukemia and stroma as well as the role of the leukemic microenvironment in drug screening studies.

Feature Scope while Logistic Regression requires pre-selected features, Dense Net is a deep learning technique that can automatically extract features from raw data. As a result, the effectiveness of both methods might be impacted by the features chosen for the comparison machine learning to white blood cells.

The volume of data used for training has a significant impact on how well deep learning algorithms function. The performance of Dense Net may not perform noticeably better than Logistic Regression if the sample size is too small.Feature Scope: While Logistic Regression requires pre-selected features, Dense Net is a deep learning technique that can

automatically extract features from raw data. As a result, the effectiveness of both methods might be impacted by the features chosen for the comparison of red blood cells.

STATISTICAL ANALYSIS

Vitamins and a genetic condition are the dependent factors, while blood tissue damage is the independent variable. The accuracy is determined using IBM SPSS software, output values are obtained using Matlab, and two parameters are compared using a sample T-test.(Garfield 2009)

RESULT

Table 1. The following table 1 consists of the comparison of accuracies of sample size 10 for densenet algorithm and Support Linear regression algorithm.

	Group	N	Mean	Std deviation	Std Error Mean
Accuracy	DN	20	85.20	2.814	0.629
Accuracy	LR	20	66.50	3.620	.809

Table 2: The following table gives group statistics for which shows a sample size of N=10 for each group. The mean percentage of the densenet algorithm is 98.61% whereas the accuracy percentage of the Linear regression algorithm is 96.31%. Standard deviation as well as standard error rate is also shown

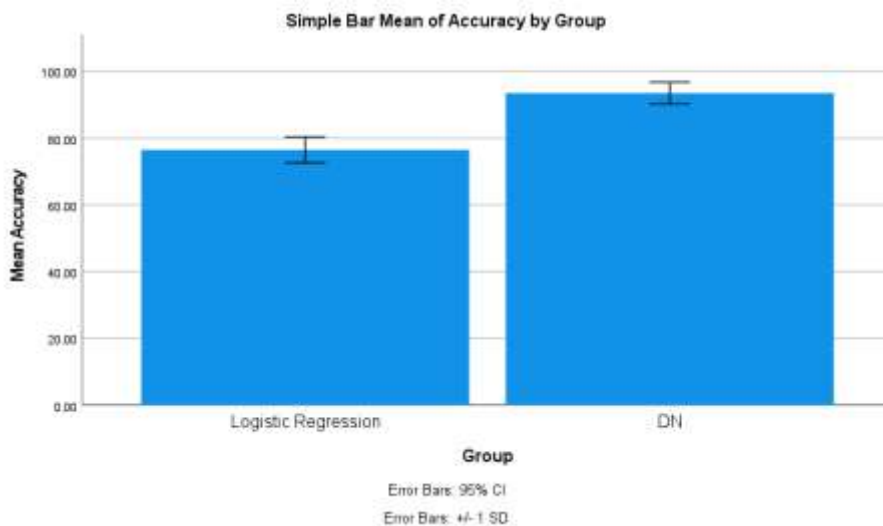
Parameter	Equal Variances	T-test for Equality of Mean								
		F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Accuracy	Equal variances assumed	3.775	.059	18.242	38	.000	18.704	1.025	16.629	20.780

	Equal Variance Not assumed			18.242	35.822	.000	18.704	1.025	16.624	20784
--	-----------------------------------	--	--	--------	--------	------	--------	-------	--------	-------

The logistic regression classifier Novel- accuracy is 89 while the densenet's accuracy is 96.5. The logistic regression classifier model was outperformed by the densenet and logistic regression models in Groups 1 and 2, respectively.

The sample data for the accuracies of the densenet algorithm and the logistic regression algorithm are shown in Table 1.

The group data for size N=10 are displayed in Table 2, together with the mean accuracies and standard deviations for the densenet and logistic regression classifier methods.



Fig

1: Simple bar graph of accuracy using densenet and Logistics Regression operators. Densenet contains high accuracy when compared to logistics regression. X-axis: Densenet vs linear regression, vs Y-axis: Accuracy +/- 1 SD

Figure 1 compares the densenet's accuracy to that The accuracy of the densenet classifier and the logistic regression classifier are compared in Fig 1. The densenet model has a greater Novel- accuracy rate than the logistic regression classification model, which has a rate of 89. In the test of independent samples, the densenet and the logistic regression classifier are highly different from one another. The significance values obtained are .000 ($p < 0.05$), which is statistically significant. Precision rates for densenet and logistic regression are shown along the X-axis

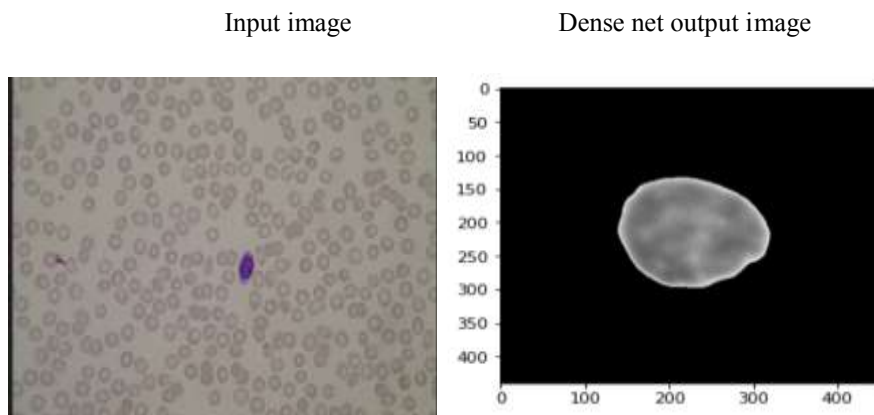


Fig. 2: Detection of leukemia disease using densenet operators in leukemia dise images.

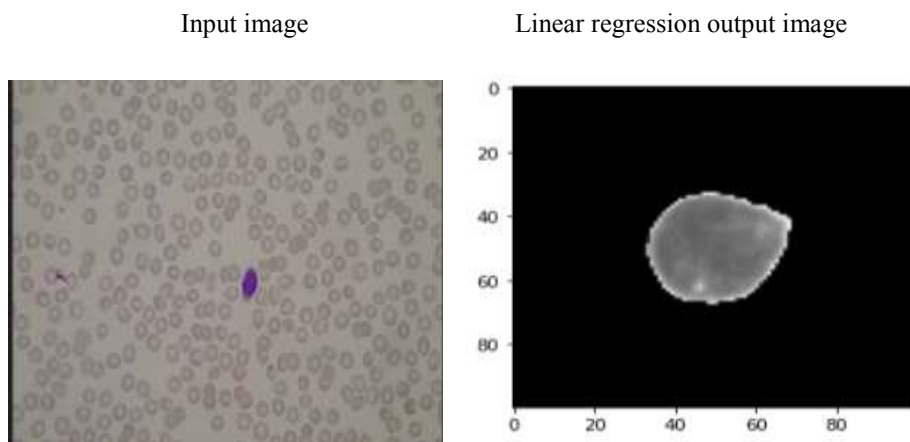


Fig 3: Detection of leukemia disease using logistics regression operators in leukemia dise images.

Fig 2 and 3 demonstrate the input and output images of Densenet and linear regression. Y-axis: Mean keyboard identification accuracy, 95% confidence interval, and 2 standard deviations.

DISCUSSION

With data and a statistical significance level of $p=.000$ ($p<0.05$), there is a statistically significant difference between dense net and logistic regression. Adults in their middle years are more likely than people under the age of 60 to develop acute promyelocytic leukemia. Children under the age of 10 do not often develop APL; nonetheless, a small number of pediatric patients under the age of 10 have developed unusual variant APL instances with Hispanics having a higher frequency of APL compared to other subtypes of AML (Lee et al. 2014). The typical NCI risk categorization of age and initial white blood cell count is insufficient to identify patients at high risk for recurrence since sickness

relapse is still the main obstacle to effective therapy. Despite the fact that the majority of patients who relapse have high risk illnesses that are clear at the time of diagnosis, a significant proportion of people with typical risks relapse and eventually pass away as a result of disease progression or treatment-related problems (Galloway, n.d.).

The evidence suggests that with each of the cytotoxic agents, the greatest threat to normal fetal development exists during the period of major organogenesis, early in gestation, even though it may be risky to draw conclusions from a small number of clinical cases and from extrapolating animal data to people (Singh et al. 2019). The extracted features are then fed into PCA to map the dataset once the dataset has been given into this model to extract its features. The input from the mapped attributes is then used to train, predict, and assess the neural network model. The features before and after mapping are supplied to the neural network for training in order to check the logic of the PCA mapping. The outcomes of the experiment are then contrasted (Garba and Chinyio 2019; Mesa, Alzate, and Varela 2017) The same experimental methodology is used concurrently with 5-fold cross-validation to compare the model performance with relevant work. This demonstrates that the test set's performance in terms of sensitivity, Novel-accuracy, and F1 is only marginally worse to the validation set. observed, the standard deviation of the test set is also lower than that of the validation set. The size of the test set is larger than the validation set, which is the cause of this (Mesa, Alzate, and Varela 2017).

In the preanalytical stage, the specific functions are focused on temperature monitoring of vehicle-based systems, which serve as intelligent transportation systems. These are tools for collecting samples at home or at a hospital that may be taken to a lab. This phase also concentrates on wireless electromagnetic sensors that are integrated into laboratory apparatus to minimize direct contact with materials during the collecting phase dense net (Sun 1992). The third item mentioned in the literature is temperature sensors. The experiment's structure is laid forth. The suggested method's strength was in its Novel-accuracy in creating the best optimal solution, according to the optimisation findings of the mathematical functions. The recommended approach was the only one, in comparison to other well-known metaheuristic optimisation algorithms including GA, DE, BA, PSO, PBA, and SOS, that could effectively solve the optimisation problem in all of the numerical functions ("An Efficient Numerical Method for the Robust Optimal Investment Problem with General Utility Functions" 2022). Due to the fact that our center is housed in a respiratory reference hospital, a limitation of the study, the majority of the cases in our study are people with lung cancer. The efficiency of the model can be increased by incorporating more examples and all recognised cancer types using artificial neural networks. While Logistic Regression requires pre-selected features, Dense Net is a deep learning technique that can automatically extract features from raw data and red blood cells. As a result, the effectiveness of both methods might be impacted by the features chosen for the comparison.

The volume of data used for training has a significant impact on how well deep learning algorithms function. The performance of Dense Net may not perform noticeably better than Logistic Regression if the sample size is too small. Feature Scope: While Logistic Regression requires pre-selected features, Dense Net is a deep learning technique that can automatically extract features from raw data. As a result, the effectiveness of both methods might be impacted by the features chosen for the comparison of white blood cells. when comparing DenseNet and logistic regression for predicting leukemia classification, some limitations to consider include the size and complexity of the data, the required computational resources, and the interpretability of the models. DenseNet may require more data and resources but can potentially provide better Novel- accuracy, while logistic regression is a simpler model with better interpretability but may not perform as well with more complex datasets.

CONCLUSION

The classification is made possible by the use of computer-based technologies for picture processing. Leukemia should be easily discernible from blood smear images thanks to the technology. We employed the Residual Convolutional Neural Network and the Dense Convolutional Neural Network (DenseNet-121) to categorize. Early and prompt discovery is crucial for choosing the best course of leukemia treatment.

DECLARATION

Conflict of interests

No conflict of interests in this manuscript.

Author Contribution

Author BV was involved in literature study, data collection, data analysis and manuscript writing. Author SJI involved in data verification, data validation and review of the manuscript.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of medical and technical Sciences(Formerly known as Saveetha University) for providing the necessary infrastructure to carry out research study.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1. HTIC Global.
2. Saveetha School of Engineering
3. Saveetha University.
4. Saveetha Institute of Medical and Technical Sciencesting. Author SJI involved in data verification, data validation and review of the Manuscript.

REFERENCES

- [1] Ahern, Holly. 2018. *Microbiology: A Laboratory Experience*.
- [2] “An Efficient Numerical Method for the Robust Optimal Investment Problem with General Utility Functions.” 2022. *Journal of Industrial and Management Optimization*. <https://doi.org/10.3934/jimo.2022212>.
- [3] Baliakas, Panagiotis, Theodoros Moysiadis, Anastasia Hadzidimitriou, Aliko Xochelli, Sabine Jeromin, Andreas Agathangelidis, Mattias Mattsson, et al. 2019. “Tailored Approaches Grounded on Immunogenetic Features for Refined Prognostication in Chronic Lymphocytic Leukemia.” *Haematologica* 104 (2): 360–69.
- [4] Brauer, Fred, Carlos Castillo-Chavez, and Zhilan Feng. 2019. *Mathematical Models in Epidemiology*. Springer Nature.
- [5] Chabner, Bruce, and Dan Louis Longo. 2018. *Cancer Chemotherapy, Immunotherapy, and Biotherapy: Principles and Practice*. LWW.
- [6] Dhara, S. 2001. “Immunohistochemical Localization of Insulin-like Growth Factors I and II at the Primary Implantation Site in the Rhesus Monkey.” *Molecular Human Reproduction*. <https://doi.org/10.1093/molehr/7.4.365>.
- [7] Dillig, Isil, and Serdar Tasiran. 2019. *Computer Aided Verification: 31st International Conference, CAV 2019, New York City, NY, USA, July 15-18, 2019, Proceedings, Part I*. Springer.
- [8] Galloway, Adele. n.d. “Adjuvant Interleukin-2, Interferon-Alpha and 5-

- Fluorouracil for Patients with High Risk of Relapse after Surgical Treatment for Renal Cell Carcinoma.” [Http://isrctn.org/](http://isrctn.org/).
<https://doi.org/10.1186/isrctn16387614>.
- [9] Garba, Etemi Joshua, and Darius Tienhua Chinyio. 2019. “COMPUTATIONAL MODEL FOR CREATING NEURAL NETWORK DATASET OF EXTRACTED FEATURES FROM IMAGES CAPTURED BY MULTIMEDIA SECURITY DEVICES.” *ICTACT Journal on Image and Video Processing*.
<https://doi.org/10.21917/ijivp.2019.0276>.
- [10] Garfield, Bob. 2009. *The Chaos Scenario*.
- [11] Lantz, Brett. 2013. *Machine Learning with R*. Packt Publishing Ltd.
- [12] Lee, Ji Heon (paul), Alfred W. Rademaker, Bayard L. Powell, Susan Geyer, Richard A. Larson, John Gregory, Nobuko Hijiyama, and James H. Feusner. 2014. “Obesity Does Not Affect the Recurrence Free Survival Rates in Children Less Than 20 Years of Age in Acute Promyelocytic Leukemia (APL).” *Blood*.
<https://doi.org/10.1182/blood.v124.21.957.957>.
- [13] Lübbert, Michael, Stefan Suciu, Liliana Baila, Björn Hans Rüter, Uwe Platzbecker, Aristoteles Giagounidis, Dominik Selleslag, et al. 2011. “Low-Dose Decitabine versus Best Supportive Care in Elderly Patients with Intermediate- or High-Risk Myelodysplastic Syndrome (MDS) Ineligible for Intensive Chemotherapy: Final Results of the Randomized Phase III Study of the European Organisation for Research and Treatment of Cancer Leukemia Group and the German MDS Study Group.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 29 (15): 1987–96.
- [14] Mesa, Fernando, Pedro Pablo Cardenas Alzate, and Carlos Alberto Rodriguez Varela. 2017. “Solution of the Problem of Cost Optimization by Using Metaheuristic Technique and Method of Optimization with Restrictions.” *Applied Mathematical Sciences*. <https://doi.org/10.12988/ams.2017.710313>.
- [15] Rodríguez, Francisco. 2021. *Models of Delay Differential Equations*.
- [16] Rozman, C., E. Montserrat, J. M. Rodríguez-Fernández, R. Ayats, T. Vallespi, R. Parody, A. Rios, D. Prados, M. Morey, and F. Gomis. 1984. “Bone Marrow Histologic Pattern--the Best Single Prognostic Parameter in Chronic Lymphocytic Leukemia: A Multivariate Survival Analysis of 329 Cases.” *Blood* 64 (3): 642–48.
- [17] Sehgal, Manav. 2017. *Data Science Solutions: Laptop Startup to Cloud Scale Data Science Workflow*.
- [18] Singh, S. P., N. Ramachandran, N. Sharma, A. K. Goel, N. M. de Sousa, J. F. Beckers, D. K. Swain, M. K. Singh, and S. D. Kharche. 2019. “Relationship of Foetal Number and Parity in Barbari Goats to Plasma Profile of Caprine Pregnancy-Associated Glycoprotein (caPAG) during Gestation and the Early Postpartum Period.” *Animal Reproduction Science*.
<https://doi.org/10.1016/j.anireprosci.2019.106190>.
- [19] Soverini, Simona, Caterina De Benedittis, Manuela Mancini, and Giovanni Martinelli. 2016. “Best Practices in Chronic Myeloid Leukemia Monitoring and Management.” *The Oncologist* 21 (5): 626–33.