ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

Selected Topological Indices with Applications to the Identification of Anti-breast Cancer Agents

Simon Mukwembi[1], Farai Nyabadza[2], Opeyemi Dorcas Ogundeji[3]

[1]University of Witwatersrand, South Africa
[2]University of Johannesburg, South Africa
[3]African Institute for Mathematical Sciences (AIMS), Cape Town, South Africa

Received: 04 January 2023; Accepted: 16 January 2023; Published: 13 February 2023

Abstract: - Breast Cancer is a global challenge that mostly affects women and its treatment is stage dependent. The treatment, among other methods, involves the use of medicines that shrink or kill the cancer cells. This type of treatment involves the use of chemical agents. In this research, the graph theoretic techniques to screen and identify chemical compounds with anti-breast cancer activities were used. In particular, the distance and graph based topological indices were used.

In this research, we developed a graph theory model that screens compounds for anti-breast cancer activity, and we also provide an assessment of the quality of model that was developed. Our newly developed graph theoretic model can predicts IC_{50} values of compounds for anti-breast cancer activity. We also recommended that normalized diameter is one of the key parameter in developing a model to predict some useful flavonoids that have anti-breast cancer activity.

We presented the upper bound on the known indices and explored their applications in the prediction of cytotoxic activity of flavonoids against breast cancer cells in drug design. A model that fits to data generated in the laboratories was formulated and analyzed. We did simulation on the correlations between parameters and IC_{50} values. Conclusions on relationship of activities and graphical structures will be made. The implications of relationship on the identification of compounds with the most activity against breast cancer are also discussed.

Keywords: Graphs ˙ Diameter ˙ Flavonoids ˙ Distances ˙ Eccentricity ˙ IC_{50} values

I. Introduction

K¨onigsberg is an ancient city of Russia, which is currently known as Kalingrad, Russia. There are seven bridges in the city which was constructed on both sides of the Pregel River. The city was divided into two islands by the seven bridges. The seven bridges are: Honey bridge, Connecting bridge, Blacksmith bridge, Green bridge, Merchant?s bridge, and Wooden bridge. Some people in the city used to spend their Sunday afternoons walking around the city. Then, the people of the city made a game for themselves, and the goal was to devise a way to walk around the city while only crossing each of the seven bridges once. No one was able to devise a path that would allow them to cross each bridge once (Natalya, 2015). The story has captivated mathematicians? imaginations up to the present day. In an attempt to solve the bridge problem some centuries ago, L. Euler introduced the concept of graphs. The solving of the bridge problem led to an increase in the use of graphs in solving real life problems. Chemical graphs were introduced in 1895 and have received increased attention in the last two decades as a result of their usefulness in chemical engineering and pharmaceuticals (Arielle, 2017). Cancer is a disease in which some cells in the body develop uncontrollably and spread to other regions of the body. Cancer can begin practically at any place in the human body, which contains billions of cells (cancer institute, 2007). There are about 100 different types of cancers. Cancers are frequently named from the organs or tissues in which they develop. Lung cancer, for example, starts in the lungs, skin cancer begins in the skin while breast cancer begins in the breast. According to Judy (2022), skin cancer is the abnormal growth of skin cells which most often develops on skin exposed to the sun. On the other hand, breast cancer is a type of cancer that develops in the cells of the breasts. Breast cancer is the second most common cancer diagnosed in women after skin cancer. Breast cancer can affect both men and women, but it mostly affects women. Substantial support for breast cancer awareness and research funding has aided in the advancement of breast cancer diagnosis and treatment. Breast cancer survival rates have increased, and the number of deaths associated with the disease have steadily decreased, and to factors such as early detection, a new personalized approach to treatment, and a better understanding of the disease. Some analgesic activity agents are used to kill pain (Pain killers). Pain in breast is one of the common symptoms that occur which makes a lot of women seek medical care. A lump in the breast, bloody discharge from the nipple and changes in the form or texture of the nipple or breast are all symptoms of breast cancer (Salah, 2015). The effectiveness and safety of a drug, in drug design, are the two main causes of drug failure. Drugs are very essential in reducing breast cancer among women. It is important to be able to identify the compounds that can produce drugs that works for breast cancer. The compounds work by shrinking or killing the cancer cells of these chemical compounds.The identification process in laboratory experiments is costly and requires a lot of experimental time. In light of these challenges, it is desirable to use mathematical models for the prediction of compound

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

which involve biological activity. The link between graph theory and chemical compounds seem to have been first presented by Cayley in 1875 when he was enumerating certain chemical compounds known as alkanes. This was followed by Harold Wiener?s attempt, in 1947, to use the Wiener index, a topological index, in predicting the boiling points of alkanes. In 1997, Sharma (1997) developed a model, based on topological indices, that could predict whether or not a compound has analgesic activity. A few such models followed (see, for instance (Mukwembi and Nyabadza, 2021b)). The model of this flavour was recently developed further by Mukwembi and Nyabadza (2021b) who presented models that could predict not only biological activity in compounds, but also the extend of the activity. Mukwembi and Nyabadza (2021b) developed a model that identifies the inhibition concentration of anti-skin cancer agents in the absence of laboratory experiments. In this thesis, diameter, Wiener index, total eccentricities will be used to determine the anti-breast cancer activity of several compounds derived from various plant sources.

Definition of terms

1.1 Graphs.

A **graph** $G=(V,E)$ consists of a finite non empty set V of elements called **vertices** and a set E of two-element subsets of V called **edges**.

The number of vertices in V is known as the **order** of G. The **size** of G is the number of edges in G.

Two vertices are said to be **adjacent** if they are joined by an edge and the vertices are called **neighbours**.

Null or empty graphs are graphs whose edge sets are empty.

A **walk** $W = u_0,u_1,u_2,\cdots,u_k$ is an alternating sequence of vertices and edges such that $\{u_i,u_{i+1}\} \in E$ for all $i = 0,1,\cdots,k-1$. A **path** is a walk in which no vertex is repeated.

We say that a graph F is a **sub graph** of a graph G if $V(F) \subseteq V(G)$ and $E(F) \subseteq E(G)$. The **distance of a subgraph** S which is denoted by $\sigma_{G(S)}$, is the sum of all distances of vertices of S to every vertex in G. Therefore,

$$\sigma_G(v) = \sum_{u \in V(G)} d_G(v,u).$$

A graph in which each pair of vertices are connected by exactly one edge is called a **complete graph**.

A graph is said to be **connected** if every pair of vertices in the graph is joined by a path. **A tree** is a connected graph with no cycles.

A **weighted graph** is a graph whose edges are labeled by numbers which are also called weights.

The **degree** of a vertex v is the number of edges incident with v. It is denoted by $deg_{(v)}$. We say that v is an **end vertex** if its degree is 1. The set of all vertices of degree greater than 1 is called **internal vertices** and the set of all end vertices in a graph is called **external vertices** (Mukwembi and Nyabadza, 2021b).

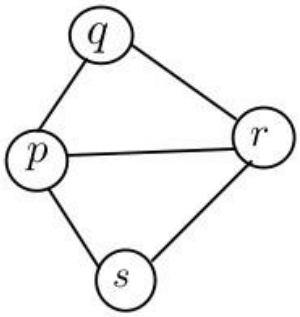


Figure 1.1: Connected graph

Figure 1.1 can be used to illustrate the notions of the degrees of a graph and its connectedness. Vertex p has degree 3, vertex q has degree 2, vertex r has degree 3, and vertex s has degree 2.

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

1.2 Concept of distances

The **distance** $_{d(u,v)}$ between vertices u and v in a graph G is the number of edges in a shortest walk joining u and v in G. Between two vertices, there may be more than one shortest walk.

For a vertex v of G, the **eccentricity** $_{ec(v)}$ of v in G is defined as the distance between v and a vertex farthest away from v. The maximum eccentricity of any vertex in a graph is known as the **diameter** of the graph. It is denoted by $diam_{(G)}$.

and so follows that $rad(G) \leq diam(G)$.

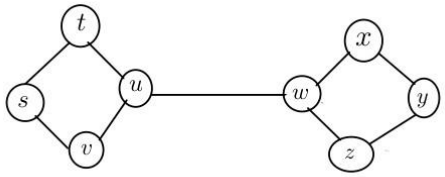


Figure 1.2: Connected graph

Figure 1.2 is used to illustrate the concepts of distance, eccentricity, radius and diameters. From Figure 1.2, the distances of the graph are:

$d(s,t) = 1$, $d(s,u) = 2$, $d(s,v) = 1$, $d(s,w) = 3$, $d(s,x) = 4$,
$d(s,z) = 4$, $d(u,y) = 3$, $d(t,y) = 4$, $d(w,y) = 2$.

From Figure 1.2, the eccentricities of the vertices of the graph are :

$ec(t) = 4$, $ec(s) = 5$, $ec(v) = 4$, $ec(u) = 3$,
$ec(x) = 4$, $ec(w) = 3$, $ec(z) = 4$, $ec(y) = 5$.

The radius of the graph in Figure 1.2 is

$$\text{rad}(G) = \min_{v \in V} ec(v) = 3.$$

The diameter of the graph in Figure 1.2 is

$$\text{diam}(G) = \max_{v \in V} ec(v) = 5.$$

1.3 Topological indices

Topological indices are graph invariants that are numerical values which are calculated on the graph to characterize the topology of the graph. It is possible to analyze mathematical values and explore various physical features of a molecule using the topological indices of the corresponding molecular graph (Islam et al., 2021). Some of the topological indices are analyzed below:

1.3.1 Wiener Index.

Consider the graph $G = (V,E)$. The **Wiener index**, $W_{(G)}$ of G, is defined as the sum of all distances between pairs of vertices of G, i.e.,

1.3.2 Graph activity.

Graph activity is a new invariant in graph theory. It was conceptualised in Mukwembi and Nyabadza (2021b) to capture cancer healing properties of the flavonoids. Let Q be the set of internal vertices and S denote the set of external vertices.

We define two invariants below,

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

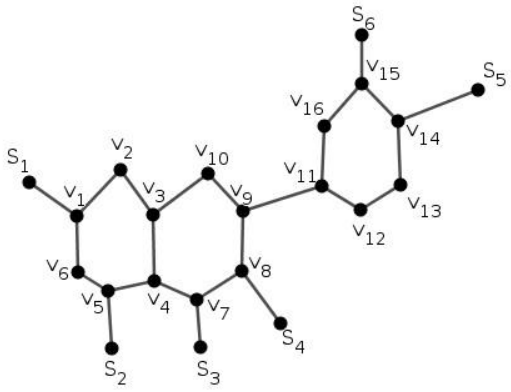


Figure 1.4: A Graph showing the molecular structure of Quercetin

Then $s(v)$, where $v = s_1$, is

$$s(v) = d(s_1,s_2) + d(s_1,s_3) + d(s_1,s_4) + d(s_1,s_5) + d(s_1,s_6).$$

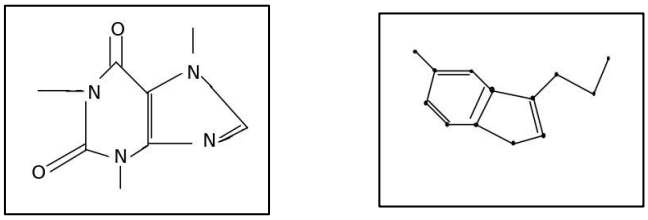
Respectively associated with each chemical compound are two numerical objects. Each compound can be represented by a mathematical object known as a graph and each compound has a mathematical numerical value linked with it which is called the Inhibition concentration value at 50%. In the next two sections, we will consider the two mathematical objects corresponding to its chemical compound separately.

1. 4 Inhibition concentration

How do we measure the extent to which a chemical compound provides a certain biological property such as healing or inhibiting certain processes? **Inhibition concentration value**, also known as IC_{50} value, is the concentration of the compound required to inhibit 50% of a biological process. It follows that a compound with a smaller IC_{50} value is more effective than the one with a bigger IC_{50} value. The possible range of the IC_{50} values should be from the lowest (small) to the highest value and it is not difficult to determine the IC_{50} values of different chemical compounds.

Molecular graph

Arthur Cayley was presumably the first to publish results that consider molecular graphs as early as 1874, indeed before the preface of the term "graph". In chemical graph theory, a molecular graph is also known as a chemical graph. The molecular graph representation is based on mapping the atoms and bonds that comprise a molecule into sets of vertices and edges. It is a representation of the structural formula of a chemical compound. Graph theory is used to mathematically model molecules in order to obtain understanding of their physical properties. Some physical properties of a substance, such as its boiling point, are connected to the geometric structure of the compound, for instance, the boiling point of alkanes is determined by the geometric structure of the alkane.



(a) Molecular structure (b) Chemical graph Figure 1.5: Representation of molecular and chemical graph

1.5.1 Flavonoids.

Flavonoids are a class of compounds that are found in plants and trees. A wide range of fruits and vegetables contain these natural compounds. Some flavonoids carry health benefits such as anti oxidation and cell signalling conditioning (Robertson, 2022).

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

Figure 1.6: Some sources of flavonoids

II. LITERATURE REVIEW

Topological indices have a wide range of applications in medical, industrial, and environmental research. Some background information and examples of applications, with a focus on diameter, total eccentricities vertices and Wiener index to develop a model that identifies anti-breast cancer agents will be presented.

According to World Health Organization (WHO), breast cancer is the second common cause of death among women (World Health Organization, 2021). As of the end of 2020, the cases of breast cancer among women who are still alive increases to 7.8 million and those who had been diagnosed with breast cancer in the previous five years, making it the most common disease in the world. In 2020, globally the number of women diagnosed with breast cancer were $2_{.3}$ million and 685000 deaths were recorded. Early-stage screening and identifying cancer types before symptoms appear to have had significant social and economic consequences (Ramirez et al., 2020). Women lose more disability-adjusted life years to breast cancer than any other kind of cancer globally. Breast cancer affects women at any age after puberty in every country throughout the world, with higher incidence in later life. The development of breast cancer begins in the living cells of the breast called epithelium (World Health Organization, 2021).

According to Poulami (2009), when cancer begins in the breast, the lumps start to grow and the changes in the breast is the important or significant ways of detecting cancer in the breast. Following the introduction of modern anticancer therapy coupled with the latest advances in science and technology in which the power of computer vision branch of artificial intelligence (T. Ige et al. 2023) had given greater insight into pattern and activities of cancerous cell in the body, the medical community was split into two camps: one claiming that isolated or synthesized chemical compounds are absolutely necessary for effective patient treatment, and the other advocating alternative cancer therapies, particularly those based on natural sources, such as plant extracts. In actuality, the two camps appear to be reconcilable: while natural sources, plant extracts, or juices can be both curative and protective, pharmaceuticals offer the best chance of inhibiting or reversing tumor formation (Levitsky and Dembitsky, 2015). Most medicinal medications are derived from plants, specifically flowers, fruits and fungi.

Although chemotherapy is the most commonly used treatment for various cancer types, acquired resistance to anticancer medications remains a critical issue, with the majority of patients who originally responded to anticancer treatments experiencing cancer recurrence (LEE et al., 2015) and the process of developing drugs is difficult because there are a lot of challenges in identifying compounds that manufacture drugs for cancer. Some of the challenges will involve extensive experimental work which requires a lot of money. Time required to complete the experiment as well as human effort are problems associated with the development of good drugs. Many plant-derived compounds have been shown to be effective in the prevention and treatment of cancer (Ahmed et al., 2016).

A much recent work in the field of pharmaceutical drug design has focused on determining chemical properties directly from their molecular structure. Both trial and error production of compounds and their random screening for activity in pharmaceutical research are time consuming and inefficient (Morgan, 2012). Mathematicians can provide models which can replace experiments which can be used to identify drug-able compounds that can be used to manufacture breast cancer drugs. Albeit, it is important to use the available cancer data to get specific patterns associated with cancer patient for intelligent automated decision (Ige & Adewale,

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

2022b) which can give a valuable prediction as well as also adopting the use of supervise, unsupervised, and self supervise machine learning model (Ige & Adewale, 2022a) for future cancer prediction on individual based on the detected pattern in cancer patients data. Topological indices can be used to establish relationships between the structure of a molecular graph and its properties or activities. A topological index, for example, can measure structural features such as size, shape, branching, symmetry, bonding patterns, and the neighborhood patterns of atoms in molecules. Topological indices have been proven useful in chemistry, bio-informatics, and network science (Qu and Cao, 2015). Key to the model that will be studied in this thesis are graph parameters which are generally known as topological indices. In the next section, each of the topological indices will be discussed separately.

The goal of graph activity is to design a model that can be used to identify drug-able compounds for skin cancer and was introduced by Mukwembi and Nyabadza (2021b). Graph activity comprises of two parameters namely; external and internal activity applied on the external vertices and internal vertices respectively, and both are treated separately. In this thesis, two parameters will be conceptualised, the first parameter aimed at the internal vertices and the other is aimed at external vertices. For the external vertices, the total Wiener index will be considered while for the internal vertices, the sum of total eccentricities and diameter will be considered as average eccentricity has itself been already introduced by Buckley and Harary (Dankelmann and Mukwembi, 2014). There it was studied as eccentric mean and research on it was initiated by AutoGraphix conjecture making computer program. Several upper bounds on the total eccentricity, a variant of average eccentricity, were presented. For instance, upper bounds in terms of order, chromatic number, domination number and connected domination number are known (Dankelmann and Mukwembi, 2014). The total eccentricities is defined as the sum of all eccentricities of vertices of the graph. In this thesis, the total eccentricities of internal vertices, will be considered and use this as one of the ingredient parameter that will help relate compounds with their IC_{50} values.

Levitsky and Dembitsky (2015) explained that the second leading cause of death for women is breast cancer. According to World Health Organization (2021), breast cancer develops in the breast tissue. It happens when breast cells mutate (change) and proliferate uncontrollably, resulting in a mass of tissue (tumor). Breast cancer, like other cancers, can penetrate and expand into the tissue that surrounds your breast. It can also spread to other places of your body and cause additional tumors to grow.

The majority of therapeutic medications are derived from plants, specifically flowers, fruits, fungi, and/or lichens. Using the chemical compounds and the experimental values of flavonoids in Ahmed et al. (2016), some of the selected anti-breast cancer agents which are found from different plant species are: Quercetin 3-O-β- D -(6" $-_n$-butyl glucuronide), Kaempferol-3-O-β- D - (6"-methyl glucuronide), Quercetin-3$-_O$-β- D -(6" -methyl glucuronide), Quercetin-3$-_O$-glucuronide (mequilianin), Kaempferol-3 $-_O$-glucuronide, Quercetin-3 $-_O$- α-rhamnopyranoside (quercitrin), Kaempferol-3-*O*glucopyranoside, Quercetin-3 -*O*-glucopyranoside (isoquercitrin), (+)-Catechin, Dehydrodicatechin A, Quercetin, Kaempferol, Adriamycin. The diagram below shows the chemical structures of the different kinds of anti-breast cancer agents.

III. RESEARCH METHODOLOGY

We developed a model that screens compounds for anti-breast cancer activity. An assessment of the quality of the model will also be presented to screen compounds for anti-skin cancer activity, we assumed a model of the form

$$\begin{aligned} IC_{50}(G) &= f([D(G)],[\zeta(G)]) \\ &= \alpha_1 + \alpha_2[D(G)] + \alpha_3[\zeta(G)] + \alpha_4[D(G)]^2 + \alpha_5[D(G)][\zeta(G)] + \alpha_6[\zeta(G)]^2 \\ &+ \alpha_7[D(G)]^2[\zeta(G)] + \alpha_8[D(G)][\zeta(G)]^2 + \alpha_9[\zeta(G)]^3, \end{aligned} \tag{1.1}$$

where $\alpha_1,\alpha_2,\cdots,\alpha_9$ are constants of model, $D(G)$ and $\zeta(G)$ are external and internal activities of the graph, respectively.

The model in (1.1) was adopted mainly because it is a Cobb-Douglas type of model for studying natural processes. The Cobb-Douglas model is generally expressed as

$$f(L,K) = Y = AL^{\alpha}K^{\beta}, \log(Y) = \log(A) + \alpha\log L + \beta\log(K),$$

where L and K are variables while α,β,A are constants.

Fitting data to the model

The Wiener index of the external vertices $W(S)$, the normalized diameter and total eccentricities of the internal vertices which is denoted by $\zeta(Q)$ will be considered as shown below :

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

$$IC_{50} = f\left(\frac{W(S)}{\binom{n+1}{3}}, \frac{\zeta(Q)}{n}\right)$$
$$= \gamma_1 + \gamma_2\frac{W(S)}{\binom{n+1}{3}} + \gamma_3\frac{\zeta(Q)}{n} + \gamma_4\left(\frac{W(S)}{\binom{n+1}{3}}\right)^2 + \gamma_5\frac{W(S)}{\binom{n+1}{3}}\frac{\zeta(Q)}{n} + \gamma_6\left(\frac{\zeta(Q)}{n}\right)^2$$
$$+ \gamma_7\left(\frac{W(S)}{\binom{n+1}{3}}\right)^2\frac{\zeta(Q)}{n} + \gamma_8\frac{W(S)}{\binom{n+1}{3}}\left(\frac{\zeta(Q)}{n}\right)^2 + \gamma_9\left(\frac{\zeta(Q)}{n}\right)^3 + \gamma_{10}\left(\frac{W(S)}{\binom{n+1}{3}}\right)^3$$
$$+ \gamma_{11}\frac{d}{n-1}\frac{\zeta(Q)}{n} + \gamma_{12}\frac{d}{n-1}\frac{W(S)}{\binom{n+1}{3}} + \gamma_{13}\left(\frac{d}{n-1}\right)^2\frac{\zeta(Q)}{n} + \gamma_{14}\left(\frac{d}{n-1}\right)^2\frac{W(S)}{\binom{n+1}{3}}, \quad (2.1)$$

where $\gamma_1, \gamma_2, \cdots, \gamma_{12}$ are constants of the model, n is the order, $W(S)$ is the Wiener index of the external vertices and $\zeta(Q)$ is the total eccentricities of the internal vertices of the graph. We now calculate the parameters for each of the graphs corresponding to the compounds in our data set.

Chem. comps	Flavonoids	Order	Diameter	$W(S)$	$\zeta(Q)$
1	Quercetin -3-O-β-D-(6"-n-butyl glucuronide)	34	11	362	217
2	Kaempferol-3-O-β-D-(6"-n-methyl glucuronide)	33	11	270	214
3	Quercetin -3-O-β-D-(6"-n-methyl glucuronide)	31	11	272	200
4	Quercetin-3-O-glucuronide	32	12	278	199
5	Kaempferol-3-O-glucuronide	33	11	282	217
6	Quercetin-3-O- α-rhamnopyranoside	38	11	282	206
7	Kaempferol-3-O-glucopyranoside	42	15	276	390
8	Quercetin-3 -O-glucopyranoside	41	14	579	328
9	Catechin	16	10	70	120
10	Dehydrodicatechin A	41	15	192	402
11	Quercetin	32	13	272	200
12	kaempferol	21	10	64	120
13	Adriamycin	35	13	332	252

Table 1: The values of $W(S)$ and $\zeta(Q)$ of the corresponding molecular graphs

IV. Data fitting

We fit data given in (2.1) and Table (1) to the model, to obtain the values of the constants in .

$\gamma_1 = -142907.95367185$, $\gamma_2 = 554530.397384071$, $\gamma_3 = 60580.6459778389$,

$\gamma_4 = 4338066.61539525$, $\gamma_5 = -308323.515737935$, $\gamma_6 = -7831.87382222808$,

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

$\gamma_7 = 469248.737469785$, $\gamma_8 = 22861.7814598866$, $\gamma_9 = 315.747150666094$,

$\gamma_{10} = -30358559.3259727$, $\gamma_{11} = 1036.42159112917$, $\gamma_{12} = 137106.433452896$,

$\gamma_{13} = 2341.90506085426$, $\gamma_{14} = -613031.22601101$.

Chem. c	Pred. IC_{50} val	Exp. val	Residue	RSS	TSS	NW(S)	N$\zeta(Q)$	N. diam
1	61.4002	61.4	-0.0002	4.6×10^{-8}	414.4776	0.0553	6.3824	0.3333
2	85.5096	85.51	0.0004	1.8×10^{-7}	1977.4101	0.0451	6.4848	0.3438
3	109.0991	109.1	0.0009	7.3×10^{-7}	4631.8445	0.0548	6.4516	0.3667
4	14.3209	14.32	-0.0009	8.0×10^{-7}	713.9902	0.0510	6,2188	0.3871
5	33.2304	33.23	-0.0004	1.8×10^{-7}	61.0127	0.0471	6,5758	0.3438
6	30.8401	30.84	-0.0001	1.2×10^{-8}	104.0681	0.0302	5.4211	0.2973
7	26.8999	26.9	2.0×10^{-5}	4.1×10^{-10}	199.9823	0.0352	9.2857	0.3659
8	38.1798	38.18	0.0002	4.5×10^{-8}	8.1893	0.0504	8.0000	0.3500
9	27.7801	27.78	-8.5×10^{-5}	7.2×10^{-9}	175.8649	0.1029	7.5000	0.6667
10	670900	67.09	-1.4×10^{-5}	1.8000	678.5256	0.0167	9.8049	0.3750
11	0.8696	0.87	0.0004	1.9×10^{-7}	1613.7843	0.0499	6.2500	0.3548
12	27.8095	27.81	0.0005	2.9×10^{-7}	175.0865	0.0416	5.7143	0.5000
13	10.5101	10.51	-0.0001	1.5×10^{-8}	932.1644	0.04650	7.2000	0.3824
	41.0415		0.0000025	11686.4006				

Table 2: The values of the diameter, normalized diameter, residual and the predicted values of the corresponding compounds in Table 1

4.1 Plotting the IC_{50} and the chemical compounds

The graph in Figure shows the correlation between the chemical compounds', experimental and predicted values. The graph is

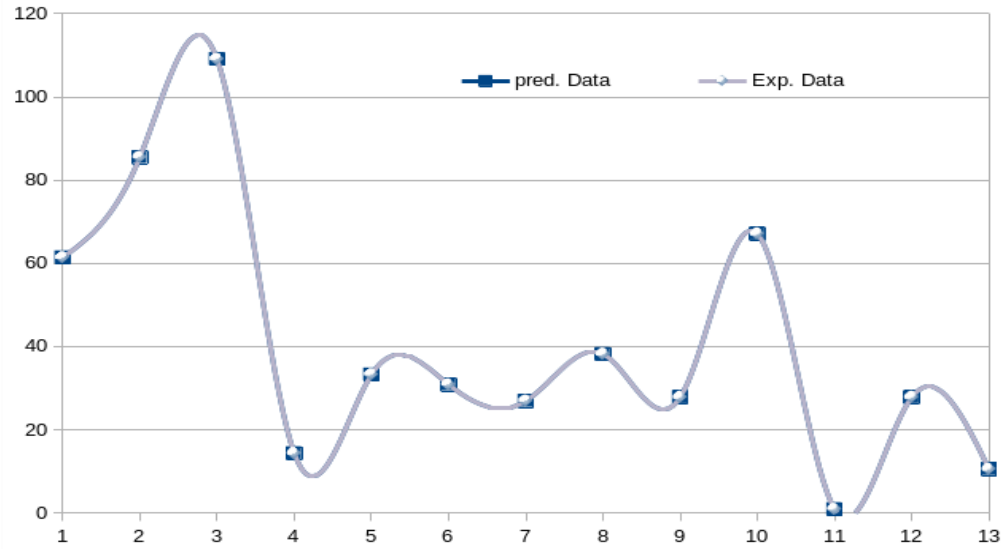


Figure 1: Graph $IC_5$0 of values

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

The coefficient of determination, CoD for the model is

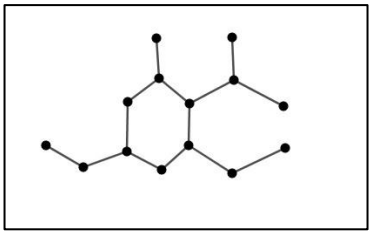
$$R^2 = 1 - \frac{RSS}{TSS},$$
$$= 1 - \frac{0.0000025066}{11686.4005936925} = 0.999999999785513,$$

The model had an accurate value of 99.999%. It accurately models the data. Therefore, the model has a goodness of fit value of 1 to 9 decimal places.

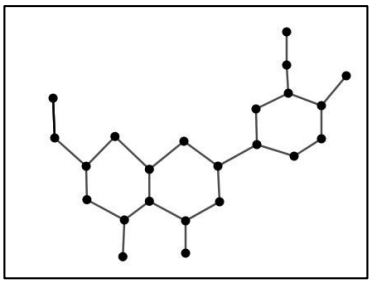
4.2 Screening and ranking of anti-breast cancer agents

Figure shown below is the graphical representations of some compounds that have not been considered.

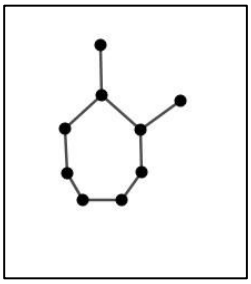
The scores of the compounds are shown in Table 3



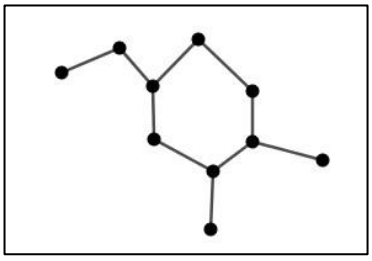
(a) Xanthoxylin graph



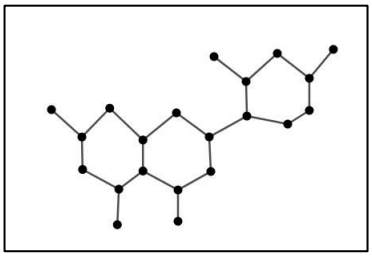
(b) Velutin graph



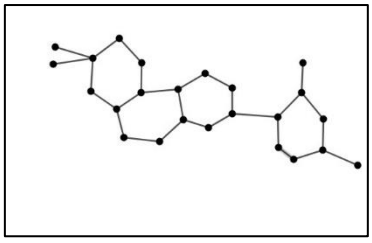
(c) Tropolone graph



(d) Kojic acid graph



(e) Norartocarpetin graph



(f) Glabridine graph

Figure 2: Graphical representations of some chemical compounds

Table 3 shows the predictions of the chemical compounds in Figure 3 and the 51 compounds in Figure 4 and their scores.

S/N	Flavonoids	IC_{50} values (Model)	NW(S)	Nζ(Q)	Nd
1	Kojic acid	-25125.2915	0.0303	2.9	0.5556
2	Tropolone	-29848.1411	0.025	2.6667	0.5
3	Norartocarpetin	-260.9655	0.0442	5.7143	0.5
4	Glabridine	3159.8649	0.0235	7.7917	0.521
5	Streppogenin	-260.9655	0.0442	5.7143	0.5
6	Dihydromorin	102.4875	0.03614	5.4545	0.4762
7	Glyasperin	486.4674	0.04538	7.16	0.5
8	Velutin	1001.1394	0.0385	6.4348	0.4545
9	Ombuin	339.8499	0.0469	6.5833	0.4783
10	Tamarixetin	-268.4174	0.0509	6.0869	0.5

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

11	Rhamnetin	-229.1148	0.0509	6	0.4545
12	Luteolin	-389.9842	0.0455	5.7143	0.5
13	Luteolin 7-methyl ether	657.3544	0.04178	6.2727	0.4762
14	5,7,3',5',-Tetrahydroxyflavanone	-1110.5659	0.0442	5.3333	0.45
15	Blumeatin	383.1267	0.0407	5.8636	0.4762
16	Dihyroquercetin -7,4'-dimethylether	339.8499	0.0469	6.5833	0.4783
17	Dihydroquercetin-4'-methylether	-268.4174	0.0509	6.0869	0.5
18	Chrysosplenol C	-183.2836	0.0526	6	0.4
19	Diosmetin	775.9514	0.0418	6.3636	0.5238
20	3,5,7-Trihydroxy-3',4'-dimethoxyflavone	160.7056	0.0469	6.2083	0.4783
21	3',4',5-Trihydroxy-3,7-dimethoxyflavone	74.2067	0.0469	6.0833	0.4348
22	3,5-Dihydroxy-3',4',7-trimethoxyflavone	548.8645	0.0435	6.72	0.4583
23	Xanthoxylin	-15218.4857	0.1099	3.1429	0.4615
24	3,3',5,7-Tetrahydroxy-4'-methoxyflavanone	-268.4174	0.0509	6.0869	0.5
25	3',5,5',7-Tetrahydroxyflavanone	-1110.5659	0.0442	5.3333	0.45
26	3,3',4',5-Tetrahydroxy-7-methoxyflavanone	-229.1148	0.0509	6	0.4545
27	3,3',5-Trihydroxy-4',7-dimethoxyflavanone	339.849	0.0470	6.5833	0.4783
28	Cyclomorusin	735.5957	0.0335	6.9677	0.3667
29	Morusin	132.8701	0.0460	6.64529	0.3667
30	Kuwanon C	151.2789	0.0589	6.3226	0.3667
31	Myricetin	-2108.4778	0.0672	5.2174	0.4545
32	Fisetin	-121.8340	0.04286	5.7143	0.5
33	Galangin	3368.2056	0.0218	5.6	0.4737
34	Gossypetin	-2112.4045	0.0662	5.2174	0.4545
35	Azaleatin	-280.5556	0.0509	5.9565	0.4545
36	Morin	-1426.6183	0.0536	5.4545	0.4762
37	Rutin	2441.7147	0.0420	9.5476	0.3902
38	Oxyresveratrol	-251.7542	0.0454	5.9444	0.5882
39	Quercetin -3-O-β-D-(6'-n-butyl glucuronide)	61.4002	0.0553	6.3824	0.3333
40	Kaempferol-3-O-β-D-(6'-n-methyl glucuronide)	85.5096	0.0451	6.4848	0.3438
41	Quercetin -3-O-β-D-(6'-n-methyl glucuronide)	109.0991	0.0548	6.4516	0.3667
42	Quercetin-3-O-glucuronide	14.3209	0.0509	6.2188	0.3871
43	Kaempferol-3-O-glucuronide	33.2304	0.0471	6.5758	0.3438

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

44	Quercetin-3-O- α-rhamnopyranoside	30.8401	0.0302	5.4211	0.2973
45	Kaempferol-3-O-glucopyranoside	26.8999	0.0352	9.2857	0.3659
46	Quercetin-3 -O-glucopyranoside	38.1798	0.0504	8	0.35
47	Catechin	27.7801	0.1029	7.5	0.6667
48	Dehydrodicatechin A	67.0900	0.0167	9.8049	0.375
49	Quercetin	0.8696	0.0499	6.25	0.3548
50	Kaempferol	27.8095	0.0416	5.7143	0.5
51	Adriamycin	10.5101	0.0465	7.2	0.3824

Table 3: The values of the Normalised Wiener index, $\zeta(Q)$ and the normalized diameter of the corresponding compounds.

4.3 Ranking

Here, the ranking of the chemical compounds in increasing order of IC_{50} values of the chemical compounds are presented to know the compound with the best IC_{50} values for treating breast cancer.

S/N	Flavonoids	Ranking IC_{50}	Order
1	Tropolone	-29848.1411172807	9
2	Kojic acid	-25125.2915201902	10
3	Xanthoxylin	-15218.4856955893	14
4	Gossypetin	-2112.40448681424	23
5	Myricetin	-2108.47782995186	23
6	Morin	-1426.6182941492	22
7	5,7,3',5',-Tetrahydroxyflavanone	-1110.56593566856	21
8	3',5,5',7-Tetrahydroxyflavanone	-1110.56593566856	21
9	Luteolin	-389.984184727185	21
10	Azaleatin	-280.555547624017	23
11	Tamarixetin	-268.417352090154	23
12	Dihydroquercetin-4'-methylether	-268.417352090154	23
13	3,3',5,7-Tetrahydroxy-4'-methoxyflavanone	-268.417352090154	23
14	Norartocarpetin	-260.965515085937	21
15	Streppogenin	-260.965515085937	21
16	Oxyresveratrol	-251.754235381437	18
17	Rhamnetin	-229.11479353	23
18	3,3',4',5-Tetrahydroxy-7-methoxyflavanone	-229.11479353	23
19	Chrysosplenol C	-183.283605259828	26
20	Fisetin	-121.833961217289	21
21	Quercetin	0.869556455999827	32
22	Adriamycin	10.5101220336282	35
23	Quercetin-3-O-glucuronide	14.3208946423383	32
24	Kaempferol-3-O-glucopyranoside	26.8999798439354	42

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

25	Catechin	27.780085074668	16
26	kaempferol	27.8094652451573	21
27	Quercetin-3-O- α-rhamnopyranoside	30.8401106899398	38
28	Kaempferol-3-O-glucuronide	33.2304285613154	33
29	Quercetin-3 -O-glucopyranoside	38.1797878740963	41
30	Quercetin -3-O-β-D-(6" -n-butyl glucuronide)	61.4002136495242	34
31	Dehydrodicatechin A	67.0900135265649	41
32	3',4',5-Trihydroxy-3,7-dimethoxyflavone	74.2066611513028	24
33	Kaempferol-3-O-β-D-(6" -n-methyl glucuronide)	85.5095704072383	33
34	Dihydromorin	102.487482783872	22
35	Quercetin -3-O-α-D-(6" -n-methyl glucuronide)	109.099145795085	31
36	Morusin	132.870794407031	31
37	Kuwanon C	151.278927237983	31
38	3,5,7-Trihydroxy-3',4'-dimethoxyflavone	160.705569210728	24
39	Ombuin	339.849924040706	24
40	Dihyroquercetin -7,4'-dimethylether	339.849924040706	24
41	3,3',5-Trihydroxy-4',7-dimethoxyflavanone	339.849924040706	24
42	Blumeatin	383.126716136759	22
43	Glyasperin	486.467449820223	25
44	3,5-Dihydroxy-3',4',7-trimethoxyflavone	548.864536342298	25
45	Luteolin 7-methyl ether	657.354407622543	22
46	Cyclomorusin	735.595714297067	31
47	Diosmetin	775.951395710683	22
48	Velutin	1001.13934999045	23
49	Rutin	2441.71471251921	42
50	Glabridine	3159.86488328851	24
51	Galangin	3368.20557317309	20

Table 4: Rankings

The research done by Ahmed et al. (2016) had quercetin as their best compound. Table 4 shows that tropolone has the best IC_{50} values amongst all screened compounds than quercetin because the model in Equation 2.1 has a goodness fit of 99.999%, the model can correctly predict the IC_{50} values of unknown compounds. According to Gusakov et al. (2021), the search for a new medicinal chemical compound for treatment of anticancer increases since World Health Organization (2021) has made it clear that breast cancer is the second leading disease among women globally. Recently, tropolone has been found to be a new pharmaceutical anti-tumor, anti-cancer and chemotherapeutic, or antimetastatic agent and the pharmacological activity in tropolone alkaloids has been detected.

V. Conclusion

We have developed a graph theoretic model that predicts IC_{50} values of compounds for anti-breast cancer activity. It is recommended that normalized diameter is one of the key parameter in developing a model to predict some useful flavonoids that have anti-breast cancer activity. The provided model can be used to screen compounds for anti-breast cancer activity. It will however be good to validate the results by laboratory experiments.

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

Our newly developed graph theoretic model can screen compounds for anti-breast cancer activity, the model can also predicts IC_{50} values of compounds for anti-breast cancer activity alongside with the recommended that normalized diameter is one of the key parameter in developing a model to predict some useful flavonoids that have anti-breast cancer activity.

References

1 Ahmed, H., Moawad, A., Owis, A., AbouZid, S., and Ahmed, O. Flavonoids of calligonum polygonoides and their cytotoxicity. Pharmaceutical biology, 54(10):2119?2126, Oct. 2016. ISSN 1388-0209, 1744- 5116. doi: 10.3109/13880209.2016.1146778. URL https://www.tandfonline.com/doi/full/10.3109/ 13880209.2016.1146778.

2 Arielle, B. Solving the k¨onigsberg bridge problem. IB Mathematics HL internal assessment, May 2017.

3 Ige, T., & Adewale, S. (2022a). Implementation of data mining on a secure cloud computing over a web API using supervised machine learning algorithm. International Journal of Advanced Computer Science and Applications, 13(5), 1–4. https://doi.org/10.14569/IJACSA.2022.0130501

4 Ige, T., & Adewale, S. (2022b). AI powered anti-cyber bullying system using machine learning algorithm of multinomial naïve Bayes and optimized linear support vector machine. International Journal of Advanced Computer Science and Applications, 13(5), 5–9. https://doi.org/10.14569/IJACSA.2022.0130502

5 Ige, T., Kolade, A., Kolade, O. (2023). Enhancing Border Security and Countering Terrorism Through Computer Vision: A Field of Artificial Intelligence. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds) Data Science and Algorithms in Systems. CoMeSySo 2022. Lecture Notes in Networks and Systems, vol 597. Springer, Cham. https://doi.org/10.1007/978-3-031-21438-7_54

6 Barefoota, L. S., R.C. Entringerbq. Extremal values for ratios of distances in trees, volume 80. Discrete Applied Mathematics, 1997.

7 Bounthavong, M. Cobb-Douglas production function and costs minimization problem. Center for innovation to implementation (Ci2i), Feb. 2019. URL https://mbounthavong.com/blog/2019/2/19/ cobbdouglas-production-function-and-total-costs. cancer institute, N. What is cancer? - national cancer institute, Sept. 2007. URL https://www.cancer. gov/about-cancer/understanding/what-is-cancer.

8 Dankelmann, C. S. S., Wayne Goddard. The average eccentricity of a graph and its subgraphs. University of Natal, Durban, 2004.

9 Dankelmann, P. and Entringer, R. Average distance, minimum degree, and spanning trees. Journal of graph theory - JGT, 33:1?13, 01 2000. doi: 10.1002/(SICI)1097-0118(200001)33:13.0.CO;2-L.

10 Dankelmann, P. and Mukwembi, S. The distance concept and distance in graphs. Distance in molecular graphs?theory, Univ. Kragujevac, Kragujevac, pages 3?48, 2012.

11 Dankelmann, P. and Mukwembi, S. Upper bounds on the average eccentricity. Discrete applied mathematics, 167:72?79, 2014. ISSN 0166-218X. doi: https://doi.org/10.1016/j.dam.2013.11.023. URL https://www.sciencedirect.com/science/article/pii/S0166218X13005738.

12 Dankelmann, P., Morgan, M., Mukwembi, S., and Swart, H. On the eccentric connectivity index and Wiener index of a graph. Quaestiones Mathematicae, 37(1):39?47, Jan. 2014. ISSN 1607-3606, 1727-

13 933X. doi: 10.2989/16073606.2013.779963. URL http://www.tandfonline.com/doi/abs/10. 2989/16073606.2013.779963.

14 Das, K. Complexity to find wiener index of some graphs. CoRR, abs/1409.1987, 2014. URL http://arxiv.org/abs/1409.1987.

15 Entringer, R. A note on spanning trees with minimum average distance, volume 17. Bull institute combination application, 1996.

16 Erdos, PACH, J., POLLACK, R., and TUZA, Z. Radius, diameter and minimum degree. B47. Journal of combinatorial theory, 1989. original-date: April 27,1989.

17 Fajtlowicz, S. On conjectures of graffiti. Department of Mathematics, University of Houston, 4800, Calhoun Road, Houston TX 77004, U.S.A., 72(1):113?118, 1988.

18 Gusakov, E. A., Topchu, I. A., Mazitova, A. M., Dorogan, I. V., Bulatov, E. R., Serebriiskii, I. G., Abramova, Z. I., Tupaeva, I. O., Demidov, O. P., Toan, D. N., Lam, T. D., Bang, D. N., Boumber, Y. A., Sayapin, Y. A., and Minkin, V. I. Design, synthesis and biological evaluation of 2-quinolyl-1,3tropolone derivatives as new anti-cancer agents. RSC Advances, 11(8):4555?4571, 2021. ISSN 2046-2069. doi: 10.1039/d0ra10610k. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8121267/. Accessed 27 May 2022.

19 Ili?, A. Eccentric connectivity index. University of Kragujevac, Kragujevac, 2011.

20 Islam, T. U., Mufti, Z. S., Ameen, A., Aslam, M. N., and Tabraiz, A. On certain aspects of topological indices. Journal of Mathematics, 2021:e9913529, May 2021. ISSN 2314-4629. doi: 10.1155/2021/ 9913529. URL https://www.hindawi.com/journals/jmath/2021/9913529/.

21 Judy, B. Breast cancer. Mayo foundation for medical education and research (MFMER), April 2022.

22 Kimberly, J. B. Chemical graph theory. Mathematics Department Indiana University of Pennsylvania, 2008.

ISSN No. 231-2705 | DOI: 10.51244/IJRSI | Volume X Issue I January 2023

23 Kouider, M. and Winkler, P. Mean distance and minimum degree. John Wiley Sons, Inc. J, 25(1):95? 99, 1997. doi: https://doi.org/10.1002/(SICI)1097-0118(199705)25:1¡95::AID-JGT7¿3.0.CO; 2-D. URL

24 LEE, Y. S., HWANG, S. G., KIM, J. K., PARK, T. H., KIM, Y. R., MYEONG, H. S., KWON, K.,

25 JANG, C. S., NOH, Y. H., and KIM, S. Y. Topological network analysis of differentially expressed genes in cancer cells with acquired gefitinib resistance. Cancer Genomics and Proteomics, 12(3):153?166, 2015. ISSN 1109-6535. URL https://cgp.iiarjournals.org/content/12/3/153.

26 Levitsky, D. O. and Dembitsky, V. M. Anti-breast cancer agents derived from plants. Natural products and bioprospecting, 5(1):1?16, Feb. 2015. ISSN 2192-2195, 2192-2209.

27 library, H. Breast cancer overview: causes, symptoms, signs, stages and types. URL https://my. clevelandclinic.org/health/diseases/3986-breast-cancer. Morgan, M. Bounds on distance-based topological indices in graphs. University of KwaZulu-Natal, 2012.

28 Morgan, M., Mukwembi, S., and Swart, H. On the eccentric connectivity index of a graph. University of Kwazulu-Natal, 311(13):1229?1234, 2011. ISSN 0012-365X. doi: https://doi.org/10.1016/j.disc. 2009.12.013. URL https://www.sciencedirect.com/science/article/pii/S0012365X09006232. Selected Papers from the 22nd British Combinatorial Conference.

29 Morgan, M. J. Bounds on distance-based topological indices in graphs. Phd thesis, University of KwazuluNatal, 2012. URL https://researchspace.ukzn.ac.za/handle/10413/10355.

30 Mukwembi, S. A note on diameter and the degree sequence of a graph. Applied mathematics letters, 25(2):175?178, 2012. ISSN 0893-9659. doi: https://doi.org/10.1016/j.aml.2011.08.010. URL https://www.sciencedirect.com/science/article/pii/S0893965911003739.

31 Mukwembi, S. and Nyabadza, F. A new model for predicting boiling points of alkanes. Scientific reports, 11, 12 2021a. doi: 10.1038/s41598-021-03541-z.

32 Mukwembi, S. and Nyabadza, F. Predicting anti-cancer activity in flavonoids - a graph theoretic approach. preprint, Oncology, Aug. 2021b. URL http://medrxiv.org/lookup/doi/10.1101/2021.08. 20.21262354.

33 Nacaro?lu, Y. and Maden, A. On the Eccentric connectivity index of unicyclic graphs. Iranian journal of mathematical chemistry, 9(1), Mar. 2018. doi:10.22052/ijmc.2017.59425.1231.

34 Natalya, N. S. C. Graph theory problems. Berkeley math circles, 2015.

35 Mathad, V. The eccentric-distance sum of some graphs. Electronic Journal of Graph Theory and Applications, 5:51?62, 04 2017. doi: 10.5614/ejgta.2017.5.1.6.

36 Qu, H. and Cao, S. On the adjacent eccentric distance sum index of graphs. Plos one, 10(6):1?12, 06 2015. doi: 10.1371/journal.pone.0129497. URL https://doi.org/10.1371/journal.pone.0129497.

37 Ramirez, R., Chiu, Y.-C., Hererra, A., Mostavi, M., Ramirez, J., Chen, Y., Huang, Y., and Jin, Y.-F. Classification of cancer types using graph convolutional neural networks. Frontiers in Physics, 8:203, June 2020. ISSN 2296-424X. doi: 10.3389/fphy.2020.00203. URL https://www.frontiersin.org/ article/10.3389/fphy.2020.00203/full.

38 Robertson, S. What-are-flavonoids. News medical life sciences, May 2022. URL https://www. newsmedical.net/health/What-are-flavonoids.aspx.

39 Salah, K. R., Awwadallah H. Pain killers. NOVA publishers international, isbn=978-1-63463-828-9, doi = 10.13140/rg.2.1.2165.3604 edition, 2015.

40 Sharma, R. Eccentric connectivity index: a novel highly discriminating topological descriptor for structureproperty and structure-activity studies, volume 37 of No. 2. Journal of chemical information and modeling, 1997. URL doi:10.1021/ci960049h. original-date: 1997.

41 Vince, A. Counting connected sets and connected partitions of a graph, volume Volume 67(2). Australasian journal of combinatorics, 2017. World Health Organization, (WHO)