VISVESVARAYA TECHNOLOGICAL UNIVERSITY
BELAGAVI - 590018

Project Report
on

"Ramification of Chaos in Production Economics"

Submitted in partial fulfilment of the requirements for the VIII Semester

Bachelor of Engineering

in

COMPUTER SCIENCE AND ENGINEERING
For the Academic Year
2019-2020
BY

Poulami Sarkar	USN 1PE16CS110
Devanand H R	USN 1PE16CS051
T Pavan	USN 1PE16CS166
Yogish	USN 1PE17CS433

UNDER THE GUIDANCE OF
Prof. Swati Sampatrao Gambhire

Assistant Professor, Dept. of CSE, PESITBSC

Department of Computer Science and Engineering
PESIT - BANGALORE SOUTH CAMPUS
Hosur Road, Bengaluru - 560100

VISVESVARAYA TECHNOLOGICAL UNIVERSITY
BELAGAVI - 590018

Project Report

on

"Ramification of Chaos in Production Economics"

Submitted in partial fulfilment of the requirements for the VIII Semester

Bachelor of Engineering

in

COMPUTER SCIENCE AND ENGINEERING
For the Academic Year
2019-2020
BY

Poulami Sarkar   USN 1PE16CS110
Devanand H R   USN 1PE16CS051
T Pavan   USN 1PE16CS166
Yogish   USN 1PE17CS433

UNDER THE GUIDANCE OF
Prof. Swati Sampatrao Gambhire
Assistant Professor, Dept. of CSE, PESIT-BSC

Department of Computer Science and Engineering
PESIT - BANGALORE SOUTH CAMPUS
Hosur Road, Bengaluru - 560100

PESIT - BANGALORE SOUTH CAMPUS
HOSUR ROAD, BENGALURU - 560100
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project work entitled "RAMIFICATION OFCHAOS IN PRODUCTION ECONOMICS " carried out by "Poulami Sarkar, Devanand H R,T Pavan, Yogish bearing USN's 1PE16CS110, 1PE16CS051, 1PE16CS166, 1PE17CS433" respectively in partial fulfillment for the award of Degree of Bachelors (Bachelors of Engineering) in Computer Science and Engineering of Visvesvaraya Technological University, Belagavi during the year 2019-2020. It is certified that all corrections/ suggestions indicated for internal assessment have been incorporated in the report. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the said Degree.

Signature of the Guide   Signature of the HOD   Signature of the Principal

Prof.Swati Gambhire   **Dr.Sandesh B. J.**   **Dr.J Surya Prasad**

Assistant Professor   **HOD, CSE**   **Principal, PESIT-BSC**

External Viva

Name of the Examiners   **Signature with Date**

1.

2.

Acknowledgements

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encourangement crowned my effort with success.

We are indebted to our Guide, Prof. Swati Sampatrao Gambhire, Assistant Professor, Department of computer Science and Engineering, PESIT-Banglore South Campus, who has not only coordinated our work but also given suggestions from time to time.

We are also extremely grateful to our Project Co-ordinators, Department of Computer Science and Engineering, PESIT Banglore South Campus, for thier constant support and advice throughout the course of preperation of this document.

We are greately thankful to Dr.Sandesh B J, Professor and HOD, Department of Computer Science and Engineering, PESIT Banglore South Campus, for his able guidance, regular source of encouragement and assistance throughout this project.

We would like to express our immense gratitude to Dr.J. Suryaprasad, Director and Principal, PESIT Banglore South Campus, for providing us with excellent infrastructure to complete our project work.

We gratefully acknowledge the help lent out to us by all faculty members of the Department of Computer Science and Engineering, PESIT Banglore South Campus, at all difficult times. We would also take this opportunity to thank our college management for the facilities provided during the course of the project. Furthermore, we acknowledge the support and feedback of my parents and friends.

POULNAMI SARKAR
DEVANAND H R
T PAVAN
YOGISH

ABSTRACT

We intend to explore the global influence (internationality) of academic institutions in terms of non-local citations(global citations). This can be of great interest to many potential students, academicians as well as researchers. The influence of institutions over time and is a non-linear system and can incorporate chaos. We posit a novel method to model this influence by using the Constant Elasticity of Substitution (CES) function optimized using the Particle swarm optimization algorithm(PSO) under chaotic conidtions.

Since such a system is non-linear,it may be subject to chaotic behaviour.We examine the chaos theoretic implications of the model via the nature of convergence of the model to a local minima and investigate if CES admits to strange attractors,fixed points etc.This is done by extracting first order ODE from the model,by approximate and regular analytical methods.

Keywords: Machine Learning,influence,CES(Constant Elasticity of Substitution Chas theory.

Contents

List of Figures

Chapter 1

Introduction

1.1   PURPOSE OF THE PROJECT

Influence of an institutions based on their research work can be quantified using the Constant Elasticity of Substitution(CES) function optimized using the Particle swarm optimization algorithm(PS0).Parameters used for this purpose are the number of citations and the number of articles published by an institution and the future influence of research done by an institution can be predicted using time series prediction.

1.2   SCOPE

The CES function was developed by Arrow, Chenery, Minhas and Solow in their new famous paper of 1961

$$Q = A[\alpha K^{\rho} + (l - \alpha)L^{\rho}]^{\frac{\eta}{\rho}}$$

CES is an economic production function that models the quantity of output produced due to a percentage change in factors K and L In our application we are assuming K and L to be the global influence of the institutions in terms of global citations and extent of international collaboration
The PSO is a method to optimise a problem by iteratively trying to improve acandidate solution It is a metaheuristic as it makes few or no assumptions about the problem being optimized and can search very large spaces of candidate solutions. It does not require the optimization problem to be differentiable, compared to classic optimization methods such as gradient descent.

1.3   PROBLEM STATEMENT

We intend to explore the scholarly influence of academic institutions which can be of great interest to many potential students,academicians as well as researches. The influence of institutions over time and is a non-linear system and can incorporate chaos. We posit a novel method to model this influence by using the Constant Elasticity of Substitution(CES) function optimized using the Particle swarm optimization algorithm(PS0) under chaotic conditions.

1.4   SUMMARY

This chapter gave a brief introduction on what exactly the proposed system is. It also covered the future scope and demand of this project. Also how this can help the other sectors. Apart from this, we are trying different approach reffering papers mentioned and trying to get accuracy as high as possible.

Chapter 2

Literature Survey

In this chapter,we see the work previously done in this area by the CES production function optimized using the Particle swarm optimization algorithm(PS0). After some of the research, following are the research papers that we have gone through in order to achieve our objective.

2.1 Strogatz Non-linear dynamics and Chao:Chaotic behviour

a. Chaotic systems are dynamical systems that are highly sensitive toinitial conditions, popularly known as butterfly effect

b. The butterfly effect causes exponential divergence of the trajectoriesof two identical chaotic systems started with nearly the same initialconditions

c. The chaos phenomenon was first observed in weather models byLorenz

d. In our application conditions such as superstar effect can cause asimilar divergence the in trajectory of an institution's influence

2.2 B. Goswami, J. Sarkar, S. Saha and S. Kar Stochastic Frontier AnalysisApproach to Revenue Modeling in Cloud Data Centers, J. CommunicationNetworks and Distributed Systems

a. The CES model has also been used for revenue optimization in cloud data centers.

b. A modified version of CES was used for optimizing cost to a minimum value

$$y = f(S, I, P, N) = (S^\rho + I^\rho + P^\rho + N^\rho)$$

c. S = cost of servers, I=cost of infrastructure, P = cost of power and N = cost of network

d. The chaos phenomenon was first observed in weather models byLorenz

e. In our application conditions such as superstar effect can cause asimilar divergence the in trajectory of an institution's influence

2.3 Suryoday Basak, Snehanshu Saha, Archana Mathur, Kakoli Bora, Simran Makhija, Margarita Safonova, Surbhi Agrawal CESSA Meets Machine Learning: From Earth Similarity to Habitability Classification of Exoplanets, Astronomy and Computings

a. The CES approach (CEESA)', has also been used to address the shortcoming of previous metrics, Cobb Douglas Habitability score.

b. Presents an example of how CES can be optimized using PSO

2.4 Arun John, Anish Murthy, Snehanshu Saha CD-HPF: Chaotic QuantumBehaved Particle Swarm Optimization for Multiobjective Optimization inHabitability Studies

a. Based on the Quantum-behaved Particle Swarm Optimization algorithm an algorithm was developed to optimize a multiobjective optimization problem, namely the Cobb Douglas Habitability function which is based on CES production functions in Economics.

2.5 Summary

This chapter explains the Literature survey for the proposed project.

Chapter 3

Project Planning

3.1 Flow Chart

1. We collect scientometric data from web of science that includes which institutions and what is the citation count.

2. Then we perform data preprocessing on data available to transform it into an understandable format by eliminating null values and data points that differs siginificantly from other observations.

3. Thirdly,we select the attributes that require for the influence of institutions over time.The attributes are Institute, Articles,Authors and Department of affiliations.

4. Then model this influence by using the Constant Elasticity ofSubstitution (CES) function optimized using the Particle swarm optimization algorithm(PSO) under chaotic conidtions.

5. $L = \frac{\sum_{p \in P(a)} \frac{Total\_number\_of\_local\_authors\_w.r.t\_institute-a\_given\_paper-p}{Total\_number\_of\_international\_authors\_w.r.t\_institute-a\_given\_paper-p}}{|P(a)|}$
   where P(a) = set of all papers with at least 1 author affiliated to institute a

   $K = \sum_{p \in P(a)} No\_of\_global\_citation$
   A citation can be considered a global citation if at least one author from the cited paper is affiliated to an international institute.

3.2   Flow Diagram
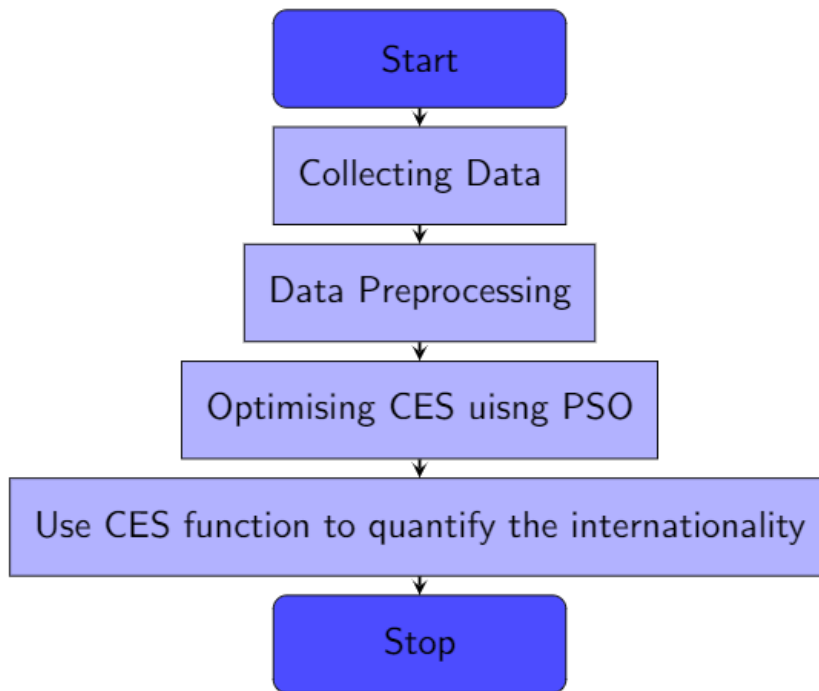
Figure 3.1: Flow diagram

Chapter 4

Hardware and Software Requirements Specification

4.1 Requirement Specification

Requirement specification is the movement of interpreting the data assembled amid investigation into prereuisite report

Software and Hardware requirements specifications are the detailed enlisting of all necessary requirements that arise in the project. The aim of having these requirements is to gain an idea of how the project isn to be implented and what is to be expected as a result of the project. The sections in this chapter deal with the various kinds of software, hardware and other functional and non functional requirements of the project.

4.2 Software Requirements

1. Python 2.7 or Python 3.5+

2. Numpy

3. Scikit-learn

4. Pandas

5. Operating system : Linux or Windows XP/7/8/10

4.3 Hardware Requirements

Hardware Requirements include:

1. GPU & CPU

2. Processor: Quad Core, 2.0 GHz or greater

3. Memory: 4 GB of RAM or more

4.4   Functional Requirements

Functional requirements are a formal way of expressing the expected servicees of a project, We have identified the unctional requirements for our project as follows:

1. The system should be able to influence of institutions over time and is a non-linearsystem and can incorporate chaos.

4.5   Non Functional Requirements

Non-functional requirements are the various capabilities offered by the system. These have nothing to do with expected results, but focus on how well the results are achieved.

1. Reliability: :The prediction subsystem should give accurate results.

2. Usability: The Overall system should be user friendly and convenient for users to operate.

3. Extensiblity: The system should be extensible i.e. it should be such that additional functional requirments can be added to the system.

4. Re-usability: The degree to which ecisting applications can be reused in new application. The predicted output could be reused in many fields.

4.6   Summary

This chapter discussed the basic software and hardware requirements. More importantly it discusses the functional and non-functional requirements.

Chapter 5

System Design

We intend to explore the global influence (internationality) of academic institutions in terms of non-local citations(global citations). This can be of great interest to many potential students, academicians as well as researchers. The influence of institutions over time and is a non-linear system and can incorporate chaos. We posit a novel method to model this influence by using the Constant Elasticity of Substitution (CES) function optimized using the Particle swarm optimization algorithm(PSO) under chaotic conditions

5.1 The CES function

In the extant literature, convex optimization principle has been used in the recent past to solvefundamental problems in science. Notwithstanding, optimization associated with technology relatedto performance in data centers and the problem of cost minimization remains an important issue. Weare aware of the cross-effects of reducing the use of one factor vis-a-vis another and apply the CESproduction function to estimate the impact of reducing the cost of the contributing factors on revenuemaximization at the data centers. It is well-known that CES belongs to the family of neoclassicalproduction functions and the CES production function for two inputs can be represented in the formbelow:

$$Q(L,K) = (\alpha K^\rho + (1-\alpha)L^\rho)^{\eta/\rho}, \tag{5.1}$$

where

Q= Quantity of output (Inteernationality score)

K = ICR (International Collaboration Ratio)

L = NLIQ (Non- Local Influence Quotient)

$\rho =$ Elasticity of substitution

and $\alpha =$ Share parameter

5.2 CES derivations

In oder to minimize the CES function we attempt to find the first order derivative of the CES equation with respect to ρ. In our research, we explore two methods for differentiating Q with respect to ρ; 1) using Tayor Series approximation 2) using Chain Rule.

5.2.1 CES differentiation using chain rule

$$R = [\alpha K^{\rho} + (1-\alpha)L^{\rho}]$$
$$Q = \gamma[\alpha K^{\rho} + (1-\alpha)L^{\rho}]^{\frac{\eta}{\rho}}$$

Applying log to the above equation on both sides of the equality

$$\log \frac{Q}{\gamma} = \frac{\eta}{\rho} \log \left([\alpha K^{\rho} + (1-\alpha)L^{\rho}]\right)$$

Differentiating the equation w.r.t ρ using chain rule we get

$$\frac{\gamma}{Q}\frac{dQ}{d\rho} = \frac{\eta}{\rho} \frac{1}{[\alpha K^{\rho} + (1-\alpha)L^{\rho}]} * (\alpha K^{\rho} log(K) + (1-\alpha)L^{\rho} log(L))$$
$$- \frac{\eta}{\rho^2} * log([\alpha K^{\rho} + (1-\alpha)L^{\rho}])$$

Replace $\gamma[\alpha K^\rho + (1-\alpha)L^\rho]$ with R, we get,

$$\frac{\gamma}{Q}\frac{dQ}{d\rho} = \frac{\eta}{\rho}[\frac{1}{R}[\alpha K^p log(K) + (1-\alpha)L^\rho log(L)] - \frac{1}{\rho}log(R)]]$$

$$\frac{\gamma}{Q}\frac{dQ}{d\rho} = \frac{\eta}{R\rho}[\alpha K^\rho log(K) + (1-\alpha)L^\rho log(L) - \frac{R}{\rho}log(R)]]$$

$$\frac{\gamma}{Q}\frac{dQ}{d\rho} = \frac{\eta}{R\rho}[log(\frac{K^{\alpha k^\rho}L^{L^\rho(1-\alpha)}}{R^{R/\rho}})]$$

$$\frac{dQ}{d\rho} = \frac{\eta Q}{R\rho\gamma}[log(\frac{K^{\alpha k^\rho}L^{L^\rho(1-\alpha)}}{R^{R/\rho}})]$$

5.2.2 CES derivation using taylor series approximation

The general form of the Constant Elasticity of Substitution (CES) production function (CESArrow) for two inputs is

$$Q(L,K) = \gamma(\alpha K^\rho + (1-\alpha)L^\rho)^{\eta/\rho},$$

where $Q = $ quantity of output/CEESA score, and L, K represent input parameters. Define $\rho = \frac{s-1}{s}$; $s = \frac{1}{1-\rho}$, $\rho > 0$. CEESA has as its limits the Cobb-Douglas production function (CDHS), i.e.

$$\lim_{\rho\to\infty} Q = \gamma K^{-\alpha}L^{\alpha-1}.$$

Proof: We can rewrite the above equation as

$$Q = \gamma(\alpha L^\rho + (1-\alpha)K^\rho)^{\eta/\rho},$$

$$\frac{1}{\gamma}Q = (\alpha K^\rho + (1-\alpha)L^\rho)^{\eta/\rho} \tag{5.2}$$

$$\frac{1}{\gamma}Q = \exp(\eta/\rho \cdot \ln[\alpha K^\rho + (1-\alpha)L^\rho]) \tag{5.3}$$

We consider first order Taylor expansion centered at zero of the term inside

the logarithm,

$$\alpha K^\rho + (1-\alpha)L^\rho = \alpha K^0 + (1-\alpha)L^0 + \alpha \left(K^0\right)^0 \cdot K^0 \cdot \ln(K)(\rho - 0)$$

$$= (1-\alpha)\left(L^0\right)^0 \cdot L^0 \cdot \ln(L)(\rho - 0) + \frac{(\rho - 0)^2}{2!} \cdot f^2(x)$$

$$= \alpha + (1-\alpha) + \alpha \cdot \rho \cdot \ln(K) + (1-\alpha) \cdot \rho \cdot \ln(L) + O(\rho^2)$$

$$= 1 + \alpha \cdot \rho \cdot \ln(K) + (1-\alpha) \cdot \rho \cdot \ln(L) + O(\rho^2).$$

$$\alpha K^\rho + (1-\alpha)L^\rho = 1 + \rho\left[\ln\left(K^\alpha \cdot L^{(1-\alpha)}\right)\right] + O(\rho^2). \tag{5.4}$$

Now, combining the equations 5.2 & 5.4, we obtain

$$\frac{1}{\gamma}Q = \left[1 + \rho\left(\ln\left(K^\alpha \cdot L^{1-\alpha}\right)\right) + O\left(\rho^2\right)\right]^{\eta/\rho}.$$

Define $\tau = \frac{1}{\rho}$; $\rho \longrightarrow 0$; $\tau \longrightarrow \infty$. Therefore,

$$\lim_{\rho \to 0} \frac{Q}{\gamma} = \lim_{\tau \to \infty} \frac{Q}{\gamma}$$

$$= \lim_{\tau \to \infty} \left(1 + \frac{1}{\tau} \cdot \left[\ln(K^\alpha \cdot L^{1-\alpha})\right] + O\left(\tau^{-2}\right)\right)^{\eta\tau}$$

$$= \lim_{\tau \to \infty} \left(1 + \frac{1}{\tau} \cdot \left[\ln(K^\alpha \cdot L^{1-\alpha})\right]\right)^{\eta\tau}$$

$$= \exp\left(\ln(K^\alpha \cdot L^{1-\alpha})\right)^\eta.$$

Consequently we can write:

$$\lim_{\rho \to 0} Q = \gamma(K^\alpha \cdot L^{1-\alpha})^\eta$$

Assuming elasticity of scale $\eta = 1$, and constant of elasticity $\gamma = 1$, we get

$$\lim_{\rho \to 0} Q = K^\alpha \cdot L^{1-\alpha}.$$

This is the CDHS formulation as mentioned in (Bora2016, Saha2018) and is used in this paper with kNN imputation.

$$Q^{\frac{\rho}{\eta}} = 1 + \rho[\ln(K^{\alpha}L^{1-\alpha})]$$

$$Q^{\rho_1} = 1 + \rho\eta\rho_1[\ln(K^{\alpha}L^{1-\alpha})]$$

$$\text{where} \frac{\rho}{n} = \rho_1$$

$$\rho = \eta\rho_1$$

$$\ln Q Q^{\rho_1} \frac{dQ}{d\rho_1} = \eta \ln(K^{\alpha}L^{1-\alpha})$$

$$\frac{1}{\eta}\frac{dQ}{d\rho_1} = Q^{-\rho}\frac{ln(K^{\alpha}L^{1-\alpha})}{lnQ}$$

$$\frac{dQ}{d\rho} = Q^{-\frac{\rho}{\eta}}\frac{\ln K^{\alpha}L^{1-\alpha}}{\ln Q} \tag{5.5}$$

Approximation of ρ and fixed point using 4th order Runge-Kutta

5.2.3   1st order ODE

$$kl = f(x_n)\Delta t$$

$$k2 = f(x_n + 0.5k_1)\Delta t$$

$$k3 = f(x_n + 0.5k_2)\Delta t$$

$$k4 = f(x_n + k_3)\Delta t$$

$$x_{n_1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

Here, we assume f(x) to be

$$f(x) = \frac{dQ}{d\rho} = Q'$$

5.3   Summary

In this chapter we preprocess scintometric data from web of science and accordingly use Particle swarm optimization algo-rithm(PSO) to quantify the influence of institutions overtime.

Chapter 6

Implementation

6.1 Dataset

Our data set consists of two variable that are used as input to the CES function. These are Non-Local Influence Quotient (NLIQ) and International Collaboration Ratio(ICR).

6.1.1 Non-Local Influence Quotient (NLIQ)

Defining and measuring internationality as a function of influence diffusion of scientific journals is an open problem. There exists no metric to rank journals based on the extent or scale of internationality. We propose Non-Local Influence Quotient (NLIQ) as one such parameter for internationality computation. NLIQ can be calculated as follows:

$$L = \frac{\sum_{p \in P(a)} No\_of\_global\_citation\_for\_paper\_p}{Total\_citations\_on\_paper\_p}$$

6.2 Dataset

Our data set consists of two variable that are used as input to the CES function. These are Non-Local Influence Quotient (NLIQ) and International Collaboration Ratio(ICR).

Institute	Articles	Articles w/ IC	Total Citations	Parent Country Citations	ICR	NLIQ	Ranking
Yale University	139	38	2698	1122	0.4474134631	0.5841363973	8
Brown University	113	32	2381	1027	0.4878838931	0.5686686266	53
Boston University	135	37	4170	1740	0.4906360187	0.5827338129	61
University of Groningen	108	31	1901	308	0.4908451531	0.8379800105	73
Humboldt University of Berlin	113	29	2930	697	0.5908436531	0.762116041	74
Emory University	62	15	1542	678	0.5554567901	0.560311284	80
University of Freiburg	107	24	2428	656	0.4646057237	0.7298187809	86
University of Zurich	92	40	1652	220	0.5716947328	0.8668280872	90
Eberhard Karls University of Tubingen	95	40	281	110	0.5074962699	0.6085409253	91
University of Basel	36	19	522	103	0.4037449775	0.8026819923	94
Dartmouth College	77	17	1552	483	0.4722488996	0.6887886598	94
University of Helsinki	115	52	2778	356	0.5642211965	0.871850252	96
Georgetown University	22	3	212	96	0.5377777778	0.5471698113	102
Uppsala University	102	38	5070	379	0.484691615	0.9252465483	102
Queen Mary University London	85	38	3018	462	0.467197437	0.8469184891	110
University of Birmingham	145	53	4034	712	0.5684712516	0.8235002479	112
University of Bern	63	19	1922	290	0.4775185428	0.8491155047	113
Case Western Reserve University	72	11	1268	566	0.4958996212	0.5536277603	119
University of Gottingen	64	25	1491	365	0.3608238379	0.7551978538	125
Maastricht University	47	19	958	140	0.515740586	0.8538622129	127
Radboud University Nijmegen	104	37	1842	342	0.4232953266	0.8143322476	128
University of Western Australia	98	38	2636	460	0.4095596417	0.8254931715	131
Durham University	91	48	2282	548	0.5371087992	0.7598597721	133
University of Vienna	98	54	1483	257	0.5835490372	0.8267026298	134
University of Cape Town	45	15	582	95	0.4701238462	0.8367697595	136
Tufts University	57	12	1055	376	0.3145422354	0.6436018957	139
Ulm University	62	23	804	221	0.4455561717	0.7251243781	141
University of Geneva	91	29	2941	300	0.3834914701	0.8979938796	144
University of Exeter	47	12	1179	263	0.4255073871	0.7769296014	146
University of Sussex	55	21	873	137	0.5006676309	0.843069874	146

Figure 6.1: Slice of the dataset

6.2.1   Non-Local Influence Quotient (NLIQ)

Defining and measuring internationality as a function of influence diffusion of scientific journals is an open problem. There exists no metric to rank journals based on the extent or scale of internationality. We propose Non-Local Influence Quotient (NLIQ) as one such parameter for internationality computation. NLIQ can be calculated as follows:

$$L = \frac{\sum_{p \in P(a)} No\_of\_global\_citation\_for\_paper\_p}{Total\_citations\_on\_paper\_p}$$

```
 1   Countries/Regions    records % of 291
 2   PEOPLES R CHINA 55  18.900
 3   CANADA   53  18.213
 4   USA 42   14.433
 5   INDIA    29   9.966
 6   IRAN     14   4.811
 7   FRANCE   13   4.467
 8   ITALY    11   3.780
 9   SOUTH KOREA 11   3.780
10   POLAND   10   3.436
11   GERMANY 9    3.093
12   JAPAN    9   3.093
13   NEW ZEALAND 9   3.093
14   MEXICO   8    2.749
15   ALGERIA 7    2.405
16   ENGLAND 7    2.405
17   TAIWAN   7    2.405
18   TURKEY   7    2.405
19   AUSTRALIA   6    2.062
20   SERBIA   6   2.062
21   BELGIUM 4    1.375
22   GREECE   4    1.375
23   INDONESIA   4    1.375
```

Country Wise Citation report for a university

	Institute Name	TotalCitations	Parent_Country_Cita	Nliq
1	Institute Name	TotalCitations	Parent_Country_Cita	Nliq
2	Indian Institute of Te	329	53	0.838905775
3	Indian Institute of Te	230	20	0.913043478
4	Indian Institute of Te	262	49	0.812977099
5	Yale University	1581	501	0.683111954
6	Auburn University	52	7	0.865384615
7	Indian Institute of Te	210	36	0.828571429
8	Columbia University	1549	505	0.673983215
9	University of Edinbur	860	51	0.940697674
10	Australian National U	1901	137	0.927932667
11	University of Tokyo	4303	1048	0.756448989
12	Johns Hopkins Univei	3047	1185	0.611092878
13	University of Pennsyl	8241	3017	0.633903652
14	University of Michiga	3916	1290	0.670582227
15	Ecole Polytechnique	5320	475	0.910714286
16	Amrita Vishwa Vidya	10	1	0.9

Calculated Nliq values with respective Institutions

6.2.2 International Collaboration Ratio(ICR)

The international Collaboration Ratio (ICR) is a novel metric that quantifies the extent to which each university collaborates with international collaborators. For each university, ICR can be broadly defined as the ratio between the weighted sum of the contributions of domestic authors and the weighted sum of the contributions of international collaborators.

6.2.3 Calculation of ICR

ICR calculation requires the following fields for each paper for each university.

- Number of authors who have authored the paper

- Number of authors belonging to different countries grouped by countries. Eg. 2 authors from India and 4 authors from USA.

- (Optional) Number of authors belonging to the same countries but different universities grouped by university. Eg. Out of 4 authors from India be 3 belong to PES and 1 belong to IISc.

```
AU Zaytsev, P
   Wolfslag, W
   Ruina, A
AF Zaytsev, Petr
   Wolfslag, Wouter
   Ruina, Andy
TI The Boundaries of Walking Stability: Viability and Controllability of
   Simple Models
C1 [Zaytsev, Petr] Univ Michigan, Dept Mech Engn, Ann Arbor, MI 48109 USA.
   [Wolfslag, Wouter] Delft Univ Technol, Dept Biomech Engn, NL-2628 CD Delft, Netherlands.
   [Ruina, Andy] Cornell Univ, Sibley Sch Mech & Aerosp Engn, Ithaca, NY 14853 USA.
```

Sample data: A single paper from Cornell University

6.2.4 ICR Algorithm

1. Calculate the weight of each article belonging to each institution as follows

$$Weight = \frac{\frac{No.\ of\ different\ collaborating\ countries\ for\ article}{Total\ no.\ of\ authors}}{No.\ of\ authors\ from\ institution\ to\ which\ article\ belongs}$$

2. Calculate ratio of normalised weights and the total number of articles from the university which are written with International Collaboration

$$ICR = \frac{\sum Normalised\ Weights}{Total\ no.\ of\ articles\ written\ with\ IC}$$

6.2.5 ICR Calculation Example

If an article has 5 authors where 2 authors from India, 3 from USA (2 from same institution in USA, 1 from different institution in USA)

- For Indian institution, weight of article will be

$$Weight, W1 = \frac{\frac{2}{5}}{2} \ and \ Normalised \ weight, NW1 = \frac{W1}{5}$$

- For institutions from USA, weight of articles will be

$$Weight, W2 = \frac{\frac{2}{5}}{2} + \frac{\frac{2}{5}}{1} \ and \ Normalised \ weight, NW2 = \frac{W2}{5}$$

- ICR calculation

$$ICR = \frac{\sum(NW1, \ NW2)}{1} = 0.16$$

6.3   Implementation part 1

Estimating internationality scores using PSO

We have used PSO to find the value of ρ at which the output,Q of the CES function is maximum. PSO is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality.

6.3.1 PSO algorithm

Algorithm 1 Particle Swarm Optimization Algorithm

Result: 1) Position and velocity of particles at which cost is optimal 2) Optimum cost

1: Initialize an array of particles with random positions [x1,x2,x3,....] and velocities [v1,v2,v3....] on d dimensions in the problem space
2: **while** stopping condition is not met **do**
3:     Evaluate the objective function, (in our case the CES function) for each particle
4:     **if** current particle's fittness is better than the previous best, 'pbest' **then**
5:         current particles best is saved as the 'pbest' and the 'pbest' location corresponds to thecurrent location in D-dimensional space.
6:     **end if**
7:     Compare Pbest of particles with each other and update the swarm global best location with the greatest fitness (gbest).
8:     Modify the velocity and position of the particles according to the following equation

$$v[i+1] = w*v[i] + c1*rand1*(pbest - x[i]) + c2*rand2*(gbest - x[i]) \qquad (6.1)$$

$$x[i+1] = x[i] + v[i+1] \qquad (6.2)$$

9: **end while**
10: **return** best position and cost.

6.3.2 PSO Python pseudocode

We have implemented our proposed model using the Particle Swarm Algorithm (PSO).

```python
# Import pyswarm package for PSO implementation
import pyswarms as ps
import math
import pandas as pd
import numpy as np


# Load dataset
df = pd.read_csv('data/icr-nliq-combined.csv')
df = df.fillna(0)


# Define constraints for rho
max_bound = 1.0 * np.ones(1)
min_bound = 0 * np.ones(1)
bounds = (min_bound, max_bound)


# Define options for PSO
#
#v[i+1] = w*v[i] + c1*rand1*(pbest-x[i]) + c2*rand2*(gbest-x[i])#
#
# Define variables c1,c2 and w for the above PSO equation
options = {'c1': 0.6, 'c2': 0.3, 'w': 0.9}


#Run PSO Iteratively for each record in the dataset
```

```
24
25  internationality = []
26  rho = []
27  K  =[]
28  L=[]
29  for  d  in  range(len(df))[:]:
30       q.setparams(K=df['ICR'][d],L=df['NLIQ'][d])
31       print(q.K,q.L)
32
33       # Set-up optimizer
34       optimizer = ps.single.GlobalBestPSO(n_particles=25, dimensions=1, options=
              options,bounds =bounds)
35       # Function f is the CES function which takes outputs the internationality
              score Q.
36       inter,r = optimizer.optimize(f, iters=100)
37       print('internationality: %f\Rho%f'%(inter,r))
38       # Plot the internationality
39       internationality.append(inter)
40       rho.append(r[0])
41
42  # Append to dataframe
43  df['internationality'] = internationality
44  df['rho'] =rho
```

6.4   Implementation part 2

CES derivation

We plot the graphs of Q(internationality scores) VS ρ and the graphs of the
Q'(derivative of Q) VS ρ. Through these plots we aim to establish that the CES
equation has a fixed point within the specified bounds i.e 0 to 1. Hence, the
equation converges and an optimal solution of the equation can be found using
an optimization algorithm such as the Particle Swarm Optimization(PSO) al-
gorithm. The internationality score is represented by the CES equation given
as:

$$Q(L,K) = \gamma (\alpha K^\rho + (1-\alpha)L^\rho)^{\eta/\rho},$$

where

Q= Quantity of output (Inteernationality score)

K = ICR (International Collaboration Ratio)

L = NLIQ (Non- Local Influence Quotient)

ρ = Elasticity of substitution

and α = Share parameter

The derivative of Q given by Q' can be written as:

By chain rule

$$Q' = \frac{\eta Q}{R\rho\gamma}[log(\frac{K^{\alpha k^\rho}L^{L^\rho(1-\alpha)}}{R^{R/\rho}})]$$

By Taylor series expansion

$$Q' = Q^{-\frac{\varrho}{\eta}}\frac{\ln K^\alpha L^{1-\alpha}}{\ln Q}$$

We observe how the value of internationality, Q and it's derivative Q' varies with elasricity ρ. ρ being the elasticity models how the output internationality score varies when the input parameters, NLIQ and ICR are kept constant. The changing value of Q and Q' with minute changes in ρ is an indicator of the chaotic nature of the CES function. In economics this is known as the superstar effect. The plots of Q VS ρ and Q' VS ρ also show if the CES function has a fixed point within the specified intervals.

1. 0-1

2. 0.001-0.1

3. 0.101-0.9

4. 0.9-0.99

5. > 1

Chapter 7

Results

7.1 Part 1: Estimation of internationality scores using PSO
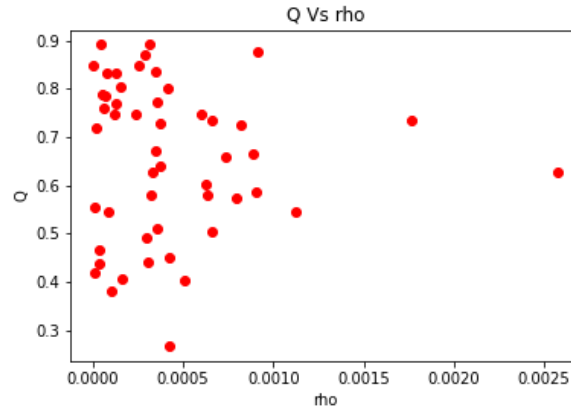
Figure 7.1: The above graph is a scatter plot of the internationality, Q VS elaticity ρ. Each point on the graph represents a separate institution.
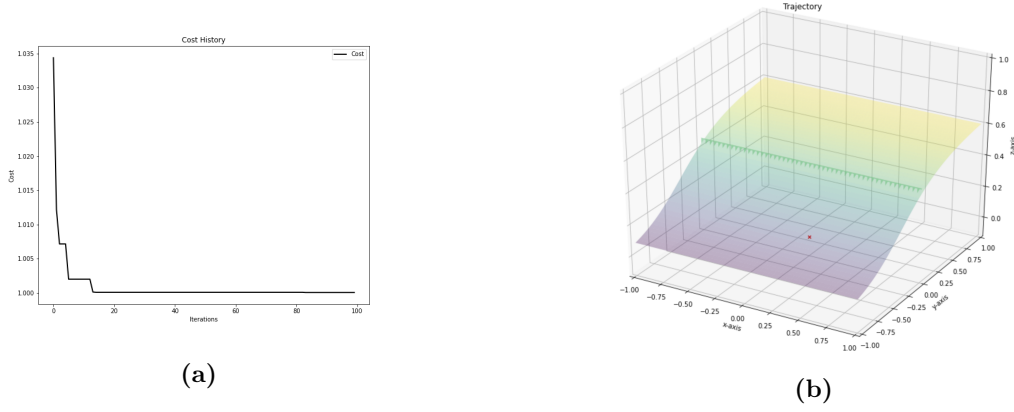
7.1.1 PSO learning curve

The below graph shows a trace of the cost history duting PSO optimization. It is the learning curve when K(ICR) = 0.01 and L(NLIQ) = 0.666666667. The values of ρ and internationality (cost) at the global optima is:

rho=0.4382448117430233 and

internationality=0.00057745

7.1.2 Result for constrained PSO optimization

We compare the different values of rho obtained after restricting the values of ρ between different intervals 1) 0.0-1.1 2) 0.001-0.1 3) 0.101-0.9 4) 0.9-0.99 5)

(a)

(b)

Figure 7.2: Learning curve for PSO optimization

> 1

This table below shows the internationality scores of each institute calculated based on the fixed NLIQ and ICR values and the optimized elasticity, ρ values. The results are arranged in the decreasing order of their internationality score.

Institution	ICR	NLIQ	0-1	0.001-0.1	0.101-0.9	0.9-0.99	¿1
University of Southern Queensland	0.75	0.90	0.8917	0.8917	0.8919	0.8930	0.0
Alexandria University	0.5833	0.9348	0.8917	0.8917	0.8926	0.8989	0.0
Auburn University	1.0	0.8653	0.8779	0.8780	0.8781	0.8788	0.9981
Australian National University	0.4929	0.9279	0.8711	0.8711	0.8726	0.8833	0.0
IISC - Bangalore	0.8571	0.8486	0.8494	0.8494	0.8494	0.8494	0.0
BITS Pilani	0.3333	0.94	0.8474	0.8475	0.8515	0.8769	0.0

Table 7.1: Cost i.e Q(Internationality) values for each institute for 6 different ranges pf ρ. This table shows the NILQ, ICR and cost values for the first 5 universities. The value of alpha=0.1

Institution	ICR	NLIQ	0-1	0.001-0.1	0.101-0.9	0.9-0.99	¿1
Auburn University	1.0	0.8653	0.9303	0.9303	0.9305	0.9324	0.9992
IISC - Bangalore	0.8571	0.8486	0.8528	0.8528	0.8528	0.8528	0.0
University of Southern Queensland	0.75	0.90	0.8257	0.8257	0.8261	0.8291	0.0
Alexandria University	0.5833	0.9348	0.7384	0.7385	0.7405	0.7570	0.0

Table 7.2: Cost i.e Q(Internationality) values for each institute for 4 different ranges pf ρ. This table shows the NILQ, ICR and cost values for the first 5 universities. The value of alpha=0.5

Table 7.1 shows the ranking of institutions after executing PSO for the following ranges of ρ; 1) 0.0-1.1 2) 0.001-0.1 3) 0.101-0.9 4) 0.9-0.99 5) > 1 Apparently, results for all conditions are consistent and it can be noticed that the ranking results are same for most of the institutes till $rho < 0.9$. This shows that the model works in harmony for all ranges of ρ between 0 and 1. Also, score is at higher side for candidates that have high NLIQ and ICR.

Evidently the model gives more weightage to lineage independent parameters while computing scores, and brings out candidates with larger NLIQ and ICR at higher positions if these values are large.

Table 7.2 shows the ranking of the institutions when α is set to 0.5. We can see that the ranking slightly changes as equal weight is given to both ICR and NLIQ. For example Auburn university which has ICR=1 and NLIQ=0.86 is ranked 1 when $\alpha = 0.5$ but when $\alpha = 0.1$, University of Southern Queensland is ranked first. This shows that when $\alpha = 0.1$, institutions with larger values of NLIQ is ranked higher, but when $\alpha = 0.5$ equal weightage is given to both ICR and NLIQ.

Institution	ICR	NLIQ	0-1	0.001-0.1	0.101-0.9	0.9-0.99	¿1
Indian Institute of Technology (IIT) - Kharagpur	0.46333	0.8129	0.7685	0.7685	0.7696	0.7772	0.0
Simon Fraser University	0.4620	0.8364	0.7882	0.7882	0.7895	0.7981	0.0

Table 7.3: Comparison between 2 universities having similar NLIQ and ICR values.

Table 7.3 shows 2 universities having comparable NLIQ and ICR values; Simon Fraser University and IIT-Kharagpur. We see that the value of rho and internationality varies although the NLIQ and ICR values are very similar. This is at par with our initial hypothesis, where we considered this system modelled using CES to be chaotic in nature.

7.2 Part 2: CES derivation

7.2.1 Plotting output of production function of different values of ρ

Traditional approaches are aimed at varying K and L input factors and monitoring the variation of output with changing input factors such as an increase in ICR or NLIQ. However, in economic models, the output of a system can also change due to unexplainable random events, popularly known as the superstar effect. The Superstar Effect refers to the change in performance caused by the presence of a highly ranked player - a superstar - in a rank-order competition. In the CES equation this super star effect is captured by the elasticity variable ρ. In the above figure we plot the rate of change of output as well as the output with respect to the elasticity, ρ. We can observe that the rate of change of output increases with changing elasticity. Moreover, the output itself also increases with increase in elasticity.
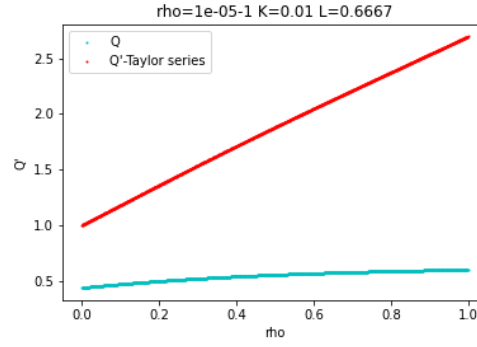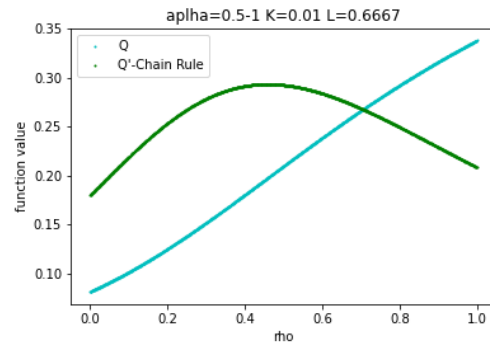
Figure 7.3: The above figure shows how the internationality (Q) varies with ρ. $\alpha = 0.1$

We are trying to find the rate of change output with respect to the an unknown factor(unexplainable value) ρ. ρ signifies an unknown factor that could cause an increase in the output Q.
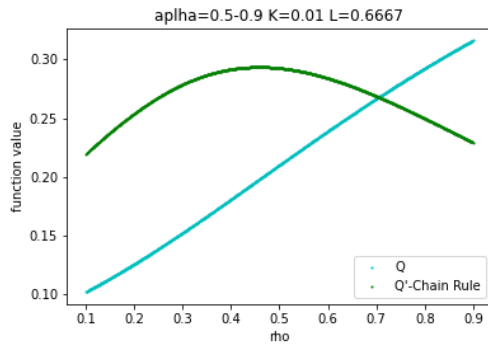
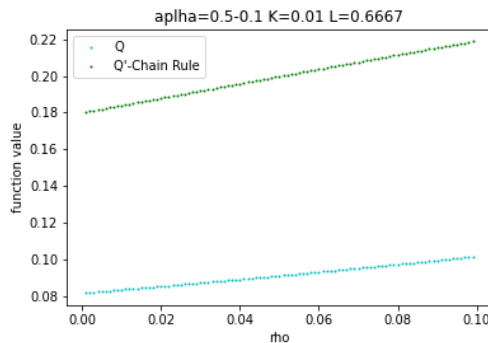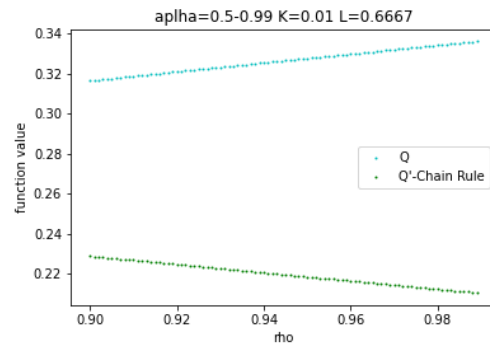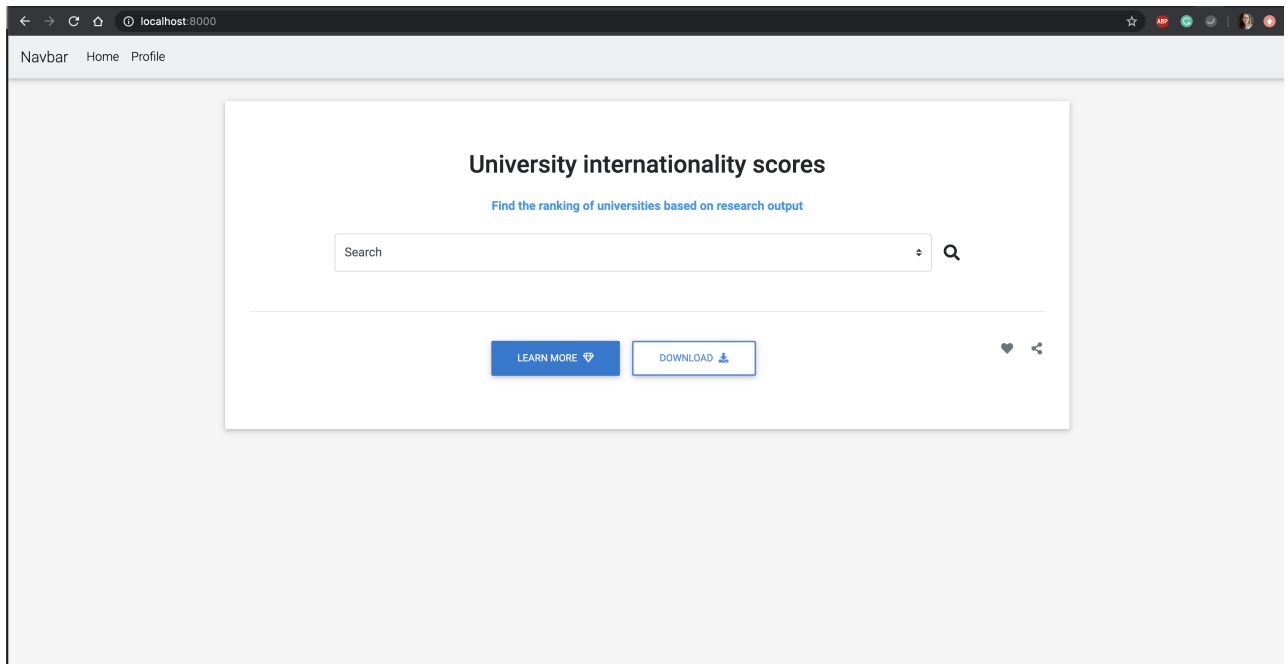**(a)** Function Value v/s $\rho(0.0 < \rho < 0.1)$

(b) [Function value v/s ρ (0.1 $< \rho <$0.99]

Figure 7.4: Plots to see how CES function changes with the values of ρ. Function v/s ρ is plotted for K=0.01 and L=0.6667. This is evident from the plots that, as value of ρ changes, it attains an optimum value.

(a) Function value v/s ρ (0.0 $< \rho <$0.1)

(b) Function value v/s ρ (0.1 $< \rho <$0.99)

Figure 7.5: Plots to see how CES function changes with the values of ρ. Function v/s ρ is plotted for K=0.01 and L=0.6667. This is evident from the plots that, as the value of ρ changes, it attains an optimum value.

.

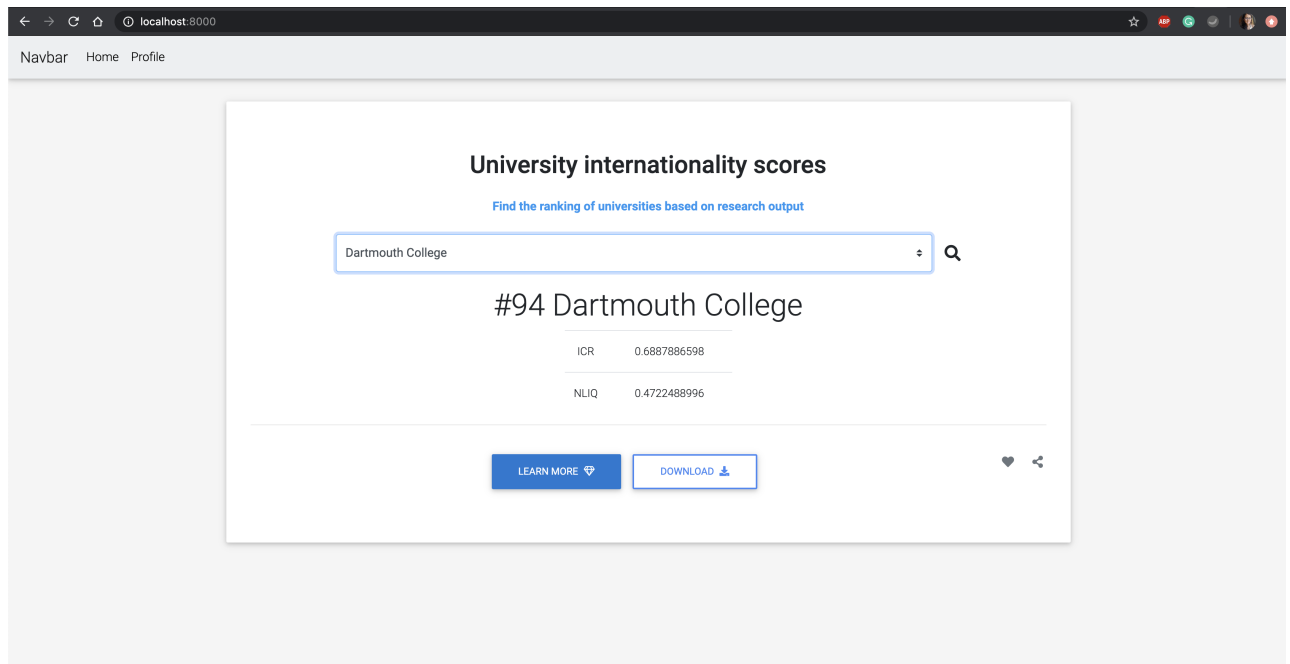## 7.3   User Interface

Figure   7.6: UI home page

Figure 7.7: Search for a university

Figure 7.8: University NLIQ, ICR and corresponding internationality scores for ρ within 5 constraints: 1) 0.0-1.1 2) 0.001-0.1 3) 0.101-0.9 4) 0.9-0.99 5) > 1
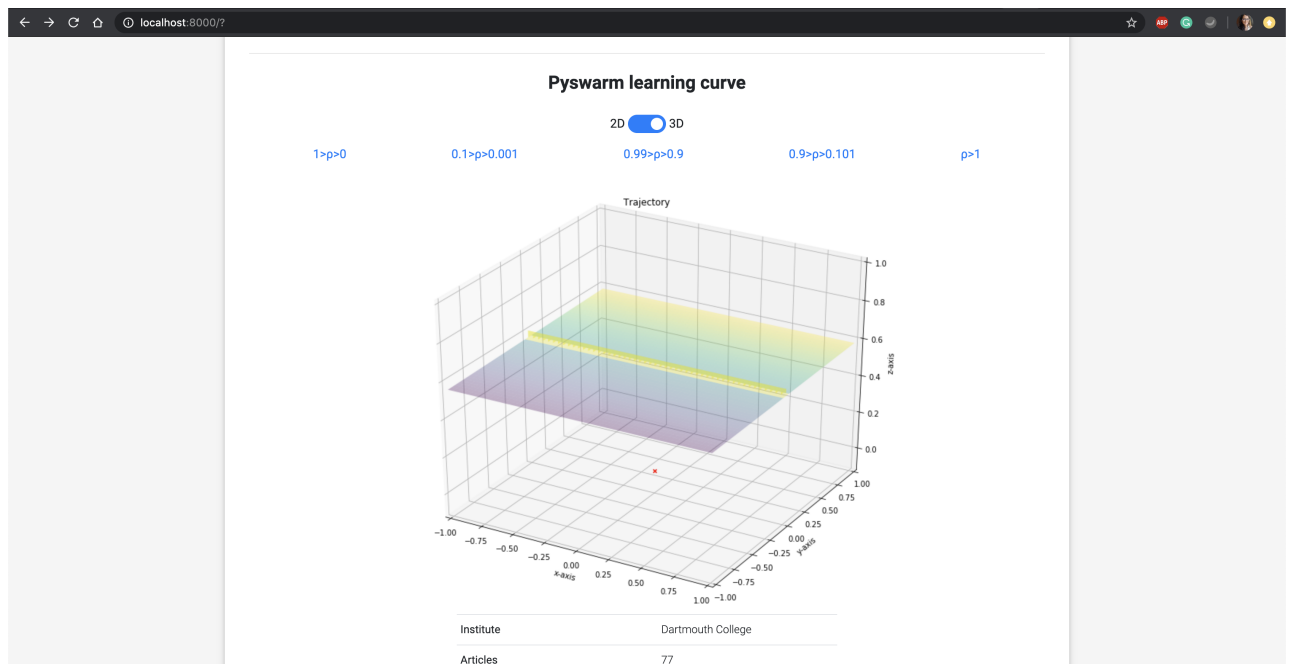
Figure 7.9: 2D Learning curve - cost history

Figure 7.10: 3D Learning curve - cost history