**IUGC**

**Adult Census Income**

**AI Project Report**

**Name: Mustansir Billah Asad Abbasi (9924)**

**Rashid Rahim (9715)**

**Ayaz Qureshi (9910)**

**Subject: Artificial Intelligence (Lab)**

**Submitted to: Sir Noman Islam**

**Date: 6th May 2017**

**Acknowledgement:**In This Project, we will predict whether income is 50000/year on less than 50000/year based on Census Data.

**Table of Content:**

|  |  |  |
| --- | --- | --- |
| **S.no** | **Heading** | **Page** |
| **1.** | **Abstract** | **3** |
| **2.** | **Introduction** | **3** |
| **3.** | **Problem Definition** | **3** |
| **4.** | **Literature Review** | **3** |
| **5.** | **Proposed Word** | **3** |
| **6.** | **Project Architecture** | **4** |
| **7.** | **Details of Works** | **4** |
| **8.** | **Experiments** | **5** |
| **9.** | **Result** | **5** |
| **10.** | **Discussion** | **6** |
| **11.** | **Conclusion** | **6** |

**Abstract: *-***

This Report Describe:

* Working of the Project?
* How this will be used in the future?
* How it is useful to the community?
* Our project calculates the accuracy of census data of people according to their salary.

**Introduction: *-***

This project ‘Adult Census Income’ calculates the accuracy of people salary, which one is less than 50,000 and which one is greater than 50,000 which depends solely upon user Inputs. In this Course, we have learnt about Artificial Intelligence so, we were Interested in developing such system.

**Problem Definition: *-***

On the bases of some information income of every individual is going to be calculated in our project this predict that a per having the specific info his or her salary should be greater than 50 or less than 50 if actual salary is increased 50 and info does not match the criteria it mean there is an extra work he or she does.

**Literature Review: *-***

This type of project has been implemented in USA to detect the salary of their citizen on the bases of their qualification we have taken this dataset from kaggle which mean this data has been implemented as a test and shown 80 % accuracy

**Proposed Work: *-***

The project is developed by three members Rashid Rahim, Mustansir Abbasi and Ayaz Qureshi three of us contributed almost equally , finding the data arranging it in organized form has been done by Ayaz Qureshi, Rashid Rahim has created the model slices the data performed training and testing, Mustansir Abbasi Tested the project wrote the user data in exel to show testing prediction and design the Django site and converted the project In dJango, other information collection and report writing has been done equally by three of us.

**Q. Why we chose this Project?**

This project was chosen because in this project, according to the Dataset, this data was extracted from the 1994 Census bureau database by Ronny Kohavi and Barry Becker (Data Mining and Visualization, Silicon Graphics). A set of reasonably clean records was extracted using the following conditions:

* ((AAGE>16) && (AGI>100) && (AFNLWGT>1) && (HRSWK>0)).
* The prediction task is to determine whether a person makes over 50000 per year.

**Project Architecture: *-***

Project is developed as an artificial intelligence neural network base (ANN).

Dataset is from kaggle and imported into project by using numpy library of python

Activation fuction relu and sigmoid applied

Optimizer adam

Slicing function for data arrangement are used

Trained the data with whole dataset and testing is performed by user input

Accuracy and result of prediction is shown in the result.

**Details of Work: *-***

*In Our Project, we take the Some Following Values From User:*

* *Age*
* *Work Class (Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked)*
* *Final Weight (continuous)*
* *Qualification (Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool)*
* *Education Class No. (continuous)*
* *Marital Status (Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse)*
* *Occupation (Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces)*
* *Relationship (Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried*
* ***Race****: White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black)*
* *Gender (Female, Male)*
* *Capital Gain (continuous)*
* *Capital Loss (continuous)*
* *HPR (Hour Per Work)*
* *Native Country (United-States, Cambodia, England, Puerto-Rico, Canada, Germany, Outlying-US(Guam-USVI-etc), India, Japan, Greece, South, China, Cuba, Iran, Honduras, Philippines, Italy, Poland, Jamaica, Vietnam, Mexico, Portugal, Ireland, France, Dominican-Republic, Laos, Ecuador, Taiwan, Haiti, Columbia, Hungary, Guatemala, Nicaragua, Scotland, Thailand, Yugoslavia, El-Salvador, Trinadad &Tobago, Peru, Hong, Holand-Netherlands)*

Due to these values, prediction will be made about the salary that can be greater than 50,000 or less than 50,000 according to user Input data*.*

**Experiments: *-***

1. Dataset is formed with entries up to 32560.
2. Inputs are given by user of a concerned applicant on python.
3. These inputs should be precise in order to return a successful output.
4. The inputs are then saved to a file in pre-built file (adult1.csv) by appending an existing data.
5. The Result or prediction about the applicant is made by using Django with the url ‘localhost:8000/adultcensusincome’.
6. The result is shown on the URL mentioned above that the Salary will be greater than or Less than 50000.

**Result: *-***

In Our Project, data which exceeds 50000 is initialized with ‘1’ and the data which less than 50000 (income) is initialize with ‘0’ in our dataset. Now, For That we will try to predict the accuracy of the income from user input. If accuracy is equal to or greater than 50% then the salary will be predicted more than 50000 or if it is below 50% then the salary will be predicted below 50000.

**Discussion: *-***

Our Project helps predict the income of people upon some certain instances:

* Reads the Dataset
* Looks to each parameter of the data
* Gets trained according to these parameters
* Takes input from end user
* Explores the probabilities of who earn more than 50000 and who is less than 50000.

**Conclusion:**

Our project ‘Adult census income’ is created really to contribute to the class of income, checking, predicting whether income exceeds 50000 or not based on census Data.