WINE QUALITY PREDICTION USING IBM WATSON MACHINE LEARNING

1. INTRODUCTION

1.1 Overview

The quality of the wine is a very important part for the consumers as well as the manufacturing industries. Industries are increasing their sales using product quality certification. Nowadays, all over the world wine is a regularly used beverage and the industries are using the certification of product quality to increases their value in the market. Previously, testing of product quality will be done at the end of the production, this is time taking process and it requires a lot of resources such as the need for various human experts for the assessment of product quality which makes this process very expensive. Every human has their own opinion about the test, so identifying the quality of the wine based on humans experts it is a challenging task. There are several features to predict the wine quality but the entire features will not be relevant for better prediction.

1.2 Purpose

To improve product quality, testing is a key element that ensures product quality. Today, different types of companies are embracing and implementing new technology to verify and assess product quality. Testing the quality of a product with human expertise is an expensive and time-consuming operation that takes time to complete. For wine quality assurance, this study investigates various machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, Ada Boost Classifier, and Gradient Boosting Classifier. These strategies automate the quality assurance process by minimising human interference and utilising accessible product attributes. The research also highlights the key characteristics that can be used to forecast the values of dependent variables. One of the major factors that can be utilised for certification is wine quality assessment, and this sort of quality certification helps to ensure wine quality in the market. Fixed acidity, volatile acidity, citric acid, residualsugar chlorides, free sulphur dioxide, total sulphur dioxide, density, pH, sulphates, and alcohol are the input variables in the red wine data set. The quality is measured on a scale of one to ten, with a greater value indicating higher wine quality.

2.LITERATURE SURVEY

2.1 Existing Problem

The significance of each feature for the wine quality prediction is not yet quantified. And in terms of performance, the current accuracy is about 67.25%. Thus, in this thesis, we considered two aspects of the problems mentioned above. The first one is the study of the importance of the features for the prediction of wine quality. The secondly, performance of the prediction model can be improved using a neural network with other ordinary classifiers.

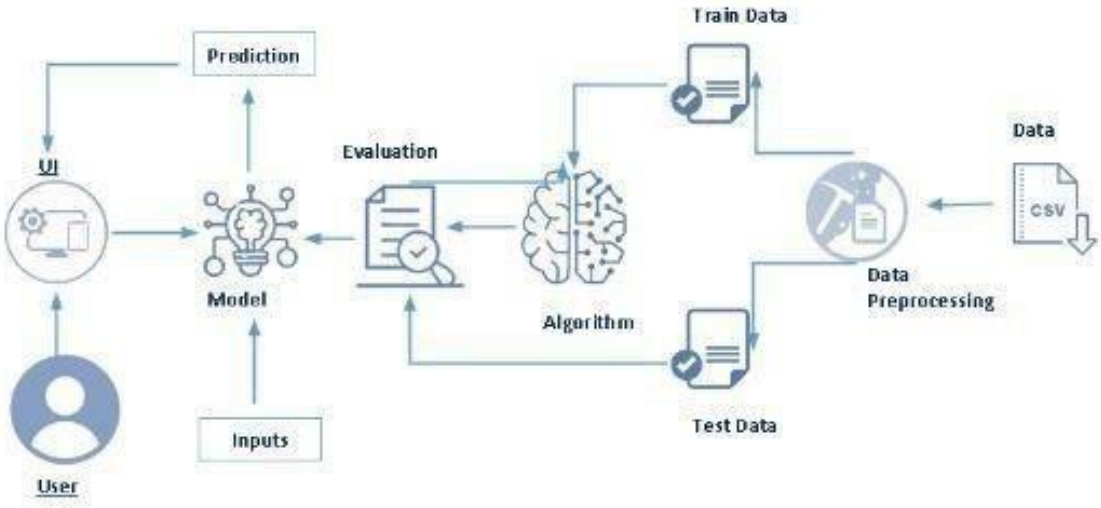
2.2 Praposed solution

The following research question and hypothesis are formulated. 1. What wine features are important to get a promising result? The researchers have used a neural network for the regression task but for the classification task neural network was never used. Hypothetically, the current prediction model that has been obtained by researchers will be improved by using the neural network. To address the research question the following objectives are formulated.

- To balance the dataset.
- To analyze the impact of the features.
- To optimize the classification models through hyperparameter tuning.
- To model and evaluate the approaches.

3.THEORITICAL ANALYSIS:

3.1 Block diagram



3.2 HARDWARE ARE AND SOFTWARE DESIGNING

Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. It was created by Guido van Rossum , and first released on February 20, 1991. Its high-level built in data structures, combined with dynamic typing and dynamic binding , make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are

available in source or binary form without charge for all major platforms, and can be freely distributed.

Anaconda Navigator

Anaconda Navigator is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications. It can be installed on Windows, Linux, and macOS.Conda is an open-source, crossplatform, package management system. Anaconda comes with so very nice tools like JupyterLab, Jupyter Notebook, QtConsole, Spyder, Glueviz, Orange, Rstudio, Visual Studio Code. For this project, we will be using Jupyter notebook and Spyder.

Jupyter Notebook

The Jupyter Notebook is an open source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter. Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself. The name, Jupyter, comes from the core supported programming languages that it supports: Julia, Python, and R. Jupyter ships with the IPython kernel, which allows you to write your programs in Python, but there are currently over 100 other kernels that you can also use.

Spyder

Spyder, the Scientific Python Development Environment, is a free integrated development environment (IDE) that is included with Anaconda. It includes editing, interactive testing, debugging, and introspection features. Initially created and developed by Pierre Raybaut in 2009, since 2012 Spyder has been maintained and continuously improved by a team of scientific Python developers and the community.Spyder is extensible with first-party and third party plugins includes support for interactive tools for data inspection and embeds Pythonspecific code. Spyder is also pre-installed in Anaconda Navigator, which is included in Anaconda.

Flask

Webframework used for building. It is a web application framework written in python which will be running in local browser with a user interface. In this application, whenever the user interacts with UI and selects emoji, it will suggest the best and top movies of that genre to the user.

Hardware Requirements: o Operating system: window 7 and above with 64bit o Processor Type Intel Core i3-3220 o RAM: 4Gb and above o Hard disk: min 100GB

4.EXPERIMENTAL INVESTIGATIONS:

4.1 Data pre-processing:

- **Removing unwanted columns.**
- **Handling missing values.**
- **Converting the target variables into binary class variables.**
- **Handling categorical data.**
- **Spilitting dataset into training and test set.**
- **Scaling Techniques.**

5.FLOW CHART:-

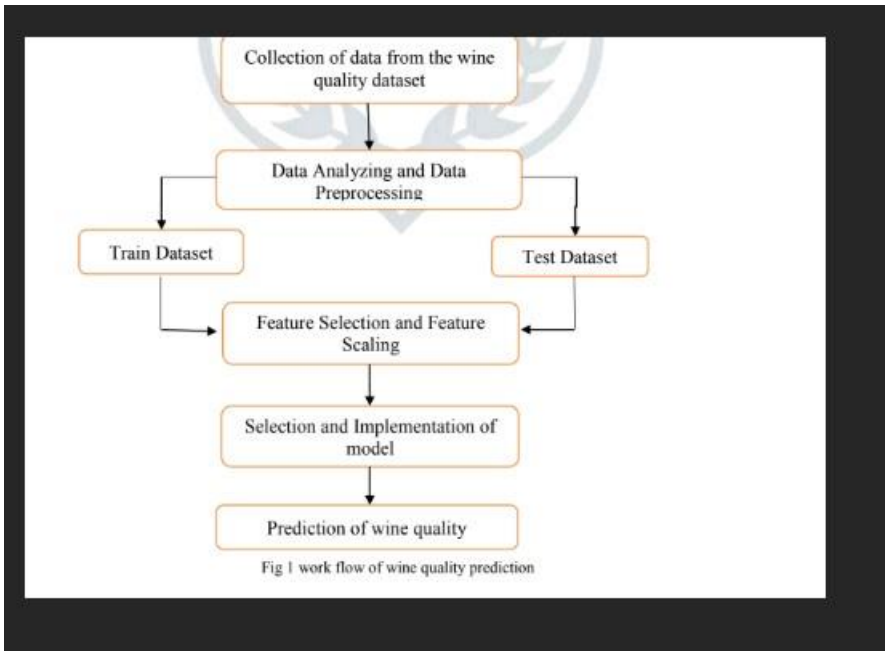
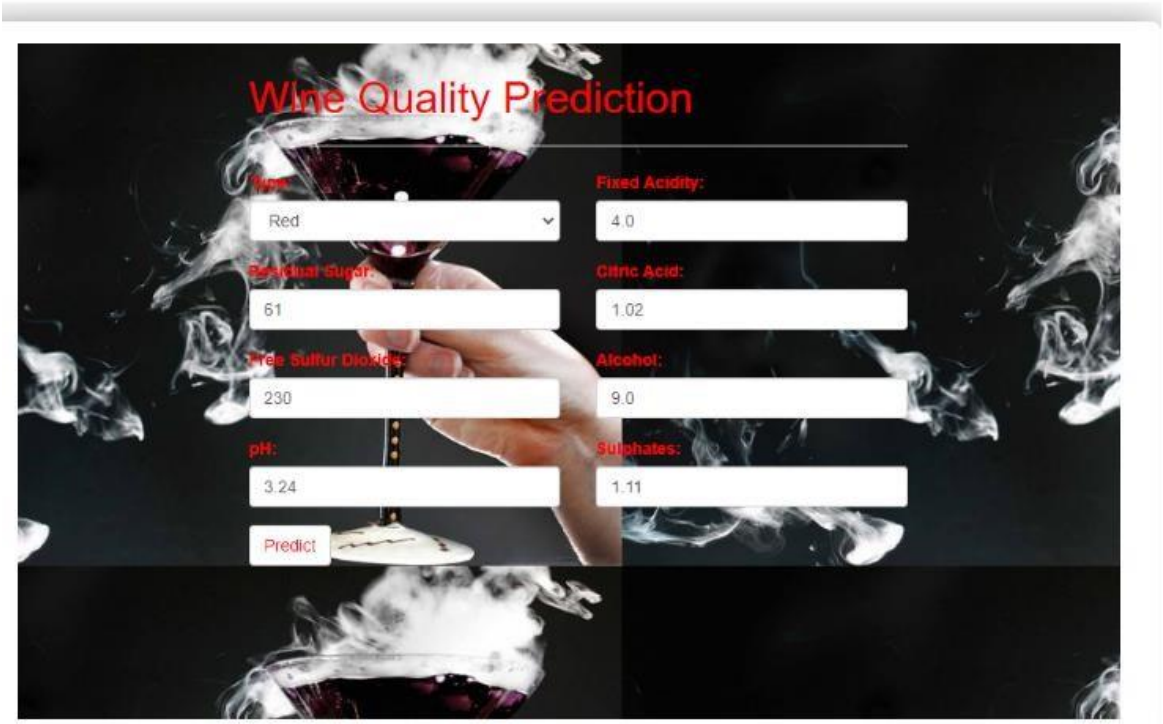


Fig 1 work flow of wine quality prediction

6. RESULTS:

Final output of the project:



7.ADVANTAGES

Easy to use

- Cost efficient
- Time efficient

8.DISADVANTAGES

● Not Applicable.

9.APPLICATIONS:-

● Results will be used by wine maufactures to improve the quality of the future wines.

● Certification bodies can also use the result for quality control.

● Results can be used to make wine selection guides for wine magazines.

● Results can be used by consumers for wine selection.

10. CONCLUSION :

In this paper, Quality of the wine is accurately predicted. Four classifier's such as Logistic Regression, Decision Tree classifier,

Random Forest Classifier and Extra Trees Classifier are used for the prediction of the quality. The contribution of this paper is

collecting the dataset of the wine and prepared that using Machine learning algorithms. The classification model is based on the 6497

records. From the analysis, we can conclude that Random Forest Classifier and Extra Trees Classifier are better with an accuracy of

88.19% and 88.79% respectively than the other methods. This model can help people with the accurate prediction of wine quality.

11.FUTURE SCOPE :

In the future, this system can be implemented further using IOT to get the real time values of the wine. In the manufacturing stage, the

sensors can be installed to collect information about the chemical components, temperature and the systems can therefore increase the

accuracy of correctness of the results. Hence, Wine quality assessment can be done in a smart way.

11.BIBILOGRAPHY:

[1] Gupta, Y. (2018), Selection of important features and predicting wine quality using

machine learning techniques, Procedia Computer Science, 125, 305–312,

doi:10.1016/j.procs.2017.12.041.

[2] Shin, T. (2020, May 8), Predicting Wine Quality with Several Classification Techniques Medium, https://towardsdatascience.com/predicting-wine-qualitywith-several-classificationtechniques-179038ea6434.

[3] Devika Pawar, Aakanksha Mahajan, Sachin Bhoithe (2019), Wine Quality Prediction using Machine Learning Algorithms, International Journal of Computer Applications Technology and Research, Volume 8–Issue 09, 385-388, ISSN:-2319–8656.

[4] Saini, A., & Sharma, A. (2019), Predicting the Unpredictable: An Application of Machine Learning Algorithms in Indian Stock Market, Annals of Data Science, doi:10.1007/s40745-019-00230-7.

[5]Zeng, Y., Liu, Y., Wu, L., Dong, H., Zhang, Y., Guo, H., Guo, Z., Wang, S.,Lan, Y. (2018). Evaluation and Analysis Model of Wine Quality Based on Mathematical Model. Studies in Engineering and Technology, 6(1), 6, doi:10.11114/set.v6i1.3626.

[6] Jambhulkar and Baporikar. (2015), Review on Prediction of Heart Disease Using Data Mining Technique with Wireless Sensor Network,International Journal of Computer Science and Applications 8 (1) 55-59.

[7] Zaveri, and Joshi. (2017), Comparative Study of Data Analysis Techniques in the domain of medicative care for Disease Predication,International Journal of Advanced Research in Computer Science

8 (3) 564-566. [8] Er, Y. (2016), The Classification of White Wine and Red Wine According to Their Physicochemical Qualities, International Journal of Intelligent Systems and Applications in Engineering, 4(Special Issue-1)

APPENDIX :- **App.py**

```
from flask import Flask, render_template, request # Flask is a application

# used to run/serve our application

# request is used to access the file which is uploaded by the user in out application

# render_template is used for rendering the html pages

import pickle # pickle is used for serializing and de-serializing Python object structures



app=Flask(__name__) # our flask app


@app.route('/') # rendering the html template def

home():
```

```
    return render_template('home.html')
@app.route('/predict') # rendering the html template def
index() :
    return render_template("index.html")

@app.route('/data_predict', methods=['GET','POST']) # route for our prediction
def predict():

    # loading model which we saved
    model = pickle.load(open('wineQuality_new.pkl', 'rb'))

    data = [[x for x in request.form.values()]]

    pred= model.predict(data)[0]
print(pred)     if pred==0:
        prediction="Bad"
else:
        prediction="Good"

    return render_template('pred.html', prediction=prediction)

if __name__ == '__main__':
app.run(debug=True)
```

home.html

```
<!DOCTYPE html>
<html>
<head>
        <title>Wine Quality Prediction Using Machine Learning</title>
```

```
        <!-- Latest compiled and minified CSS -->
        <link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.7/css/bootstrap.min.css">


        <style type="text/css">
body {
  background-image: url('https://w0.peakpx.com/wallpaper/880/804/HD-wallpaper-goodwine-red-glass-nice-drop-wine-good-beautiful-abstract.jpg')
        </style>
</head>
<body>


<nav class="navbar">
  <div class="container-fluid">
    <div class="navbar-header">
      <a class="navbar-brand" style="color:white;font-size:250%;">Wine Quality
Prediction</a>
    </div>
    <ul class="nav navbar-nav navbar-right">
        <li><a href="#" style="color:white;font-size:250%;">Home</a></li>
      <li><a href="/predict" style="color:white;font-size:250%;">Predict</a></li>
    </ul>
  </div>
</nav>


<div class="container">
  <h3 style="font-size:200%;color:#F96E6E;">Introduction</h3>
  <p style="color:white;font-size:120%;">Wine is the most healthful and most hygienic of
beverages.
  Yes, if you think deep down then you just notice that we are discussing wine, above quote
seems to be
```

right because all over the world wine was soo popular among people, and 5% of the population doesn’t

know what is wine? sounds good.

We definitely came across the fruit graphs, which is soo sweet on the test but graphs are not just to

eat, they are used to make different types of things. Wine is one of them Wine is an alcoholic drink

that is made up of fermented grapes. If you have come across wine then you will notice that wine has

also their type they are red and white wine this was because of different varieties of graphs.

You are shocked to hear that the worldwide distribution of wine is 31 million tonnes which were huge in number.

<h3 style="font-size:150%;color:#F96E6E;">What if you think about the quality of wine, how can you

differentiate the wine according to their quality? The big question arises.</h3>

 According to experts, the wine is differentiated

according to its smell, flavor, and color, but we are not a wine expert to say that wine is

good or bad. What will we do then? Here’s the use of Machine Learning comes, yes you are thinking

to write we are using machine learning to check wine quality.

</body>

</html>

Index.html

<!DOCTYPE html>

<html>

```
<head>
	<title>GDP Analysis</title>
	<!-- Latest compiled and minified CSS -->
	<link rel="stylesheet"
href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.7/css/bootstrap.min.css">
	<style type="text/css">

body {
  background-image: url('https://res.allmacwallpaper.com/get/macbook-air-
wallpapers/NameYour-Poison/7799-720.jpg');
}

	</style>
</head>
<body>

	<div class="container">
		<div class="row">
			<div class="col-md-3"></div>
			<div class="col-md-6">
				 <div class="page-header">
				<h1 style="color:red;">Wine Quality Prediction</h1>
				</div>
			</div>
		</div>
	</div>

	<div class="container">
		<div class="row">
			<div class="col-md-3"></div>
			<div class="col-md-6">
```

```
<form action="/data_predict" method="POST">
    <div class="row">
        <div class="col-md-6">
            <div class="form-group">
                <label for="type" style="color:red;">Type:</label>
                <select name="type" class="formcontrol" id="type">
                    <option value="0">Red</option>
                    <option value="1">White</option>
                </select>
            </div>
        </div>

        <div class="col-md-6">
            <div class="form-group">
                <label for="fixed_acidity" style="color:red;">Fixed Acidity:</label>
                <input type="text" class="form-control" id="fixed_acidity" name="fixed_acidity" placeholder="Range 3.8 to 15.9">
            </div>
        </div>
    </div>

    <div class="row">
        <div class="col-md-6">
            <div class="form-group">
                <label for="residual_sugar" style="color:red;">Residual Sugar:</label>
                <input type="text" class="form-control" id="residual_sugar" name="residual_sugar" placeholder="Range 0.06 to 65.8">
            </div>
        </div>
        <div class="col-md-6">
```

```
                                        <div class="form-group">
                <label for="citric_acid" style="color:red;">Citric Acid:</label>
                                            <input type="text" class="form-control" id="citric_acid" name="citric_acid" placeholder="Range 0.0 to 1.66">
                                        </div>
                                    </div>
                                </div>


                                <div class="row">
                                    <div class="col-md-6">
                                        <div class="form-group">
                                            <label for="free_sulfur_dioxide" style="color:red;">Free Sulfur Dioxide:</label>
                                            <input type="text" class="form-control" id="free_sulfur_dioxide" name="free_sulfur_dioxide" placeholder="Range (1.0, 289.0)">
                                        </div>
                                    </div>


                                    <div class="col-md-6">
                                        <div class="form-group">
                                            <label for="alcohol" style="color:red;">Alcohol:</label>
                                            <input type="text" class="form-control" id="alcohol" name="alcohol" placeholder="Range (8.0, 14.9)">
                                        </div>
                                    </div>
                                </div>
                                <div class="row">
                                    <div class="col-md-6">
```

```
                                                    <div class="form-group">
                                                        <label for="pH"
style="color:red;">pH:</label>
                                                        <input type="text" class="form-
control" id="pH" name="pH" placeholder="Range (2.72, 4.01)">
                                                    </div>
                                                </div>

                                                <div class="col-md-6">
                                                    <div class="form-group">
                                                        <label for="sulphates"
style="color:red;">Sulphates:</label>
                                                        <input type="text" class="form-
control" id="sulphates" name="sulphates" placeholder="Range (0.22, 2.0)">
                                                    </div>
                                                </div>
                                                </div>

    <button type="submit" class="btn btn-default" style="color:red;">Predict</button>
                                    </form>
                                </div>
                            </div>
                        </div>
```

Pred.html

```
<!DOCTYPE html>
<html>
<head>
        <title>Wine Quality Prediction Using Machine Learning</title>
        <!-- Latest compiled and minified CSS -->
        <link rel="stylesheet"
```

```
href="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.7/css/bootstrap.min.css">

        <style type="text/css"> body
{
  background-image: url('https://m.economictimes.com/thumb/msid-
70001605,width1200,height-900,resizemode-4,imgsize-541593/pretty-wine-
bottl.jpg'); }
        </style>
</head>
<body>
        <div class="container">
                <div class="row">
                        <div class="col-md-3"></div>
                        <div class="col-md-6">
                                 <div class="page-header">
                                <h1 style="font-size:500%;color:red;">Wine Quality
Prediction</h1>
                                </div>
                        </div>
                </div>
        </div>
        <div class="container">
                <div class="row">
                        <div class="col-md-3"></div>
                        <div class="col-md-6">
                                <div class="p-2 my-2 border">
      <h3 style="font-size:300%;color:red;">Wine Quality is predicted to be :-
{{prediction}}</h3>
                                </div>
                        </div>
                </div>
        </div>
```

<!-- Latest compiled and minified JavaScript -->

<script src="https://maxcdn.bootstrapcdn.com/bootstrap/3.3.7/js/bootstrap.min.js"></script

</body>

</html>

Style.css

@import url(https://fonts.googleapis.com/css?family=Open+Sans);

.btn { display: inline-block; *display: inline; *zoom: 1; padding: 4px 10px 4px; marginbottom: 0; font-size: 13px; line-height: 18px; color: #333333; text-align: center;text-shadow: 0 1px 1px rgba(255, 255, 255, 0.75); vertical-align: middle; background-color: #f5f5f5; background-image: -moz-linear-gradient(top, #ffffff, #e6e6e6); background-image: -mslinear-gradient(top, #ffffff, #e6e6e6); background-image: -webkit-gradient(linear, 0 0, 0 100%, from(#ffffff), to(#e6e6e6)); background-image: -webkit-linear-gradient(top, #ffffff, #e6e6e6); background-image: -o-linear-gradient(top, #ffffff, #e6e6e6); background-image: linear-gradient(top, #ffffff, #e6e6e6); background-repeat: repeat-x; filter: progid:dximagetransform.microsoft.gradient(startColorstr=#ffffff, endColorstr=#e6e6e6, GradientType=0); border-color: #e6e6e6 #e6e6e6 #e6e6e6; border-color: rgba(0, 0, 0, 0.1) rgba(0, 0, 0, 0.1) rgba(0, 0, 0, 0.25); border: 1px solid #e6e6e6; -webkit-border-radius: 4px; moz-border-radius: 4px; border-radius: 4px; -webkit-box-shadow: inset 0 1px 0 rgba(255, 255, 255, 0.2), 0 1px 2px rgba(0, 0, 0, 0.05); -moz-box-shadow: inset 0 1px 0 rgba(255, 255, 255, 0.2), 0 1px 2px rgba(0, 0, 0, 0.05); box-shadow: inset 0 1px 0 rgba(255, 255, 255, 0.2), 0 1px 2px rgba(0, 0, 0, 0.05); cursor: pointer; *margin-left: .3em; }

.btn:hover, .btn:active, .btn.active, .btn.disabled, .btn[disabled] { background-color: #e6e6e6; }

.btn-large { padding: 9px 14px; font-size: 15px; line-height: normal; -webkit-border-radius: 5px; -moz-border-radius: 5px; border-radius: 5px; }

.btn:hover { color: #333333; text-decoration: none; background-color: #e6e6e6; backgroundposition: 0 -15px; -webkit-transition: background-position 0.1s linear; -moz-transition: background-position 0.1s linear; -ms-transition: background-position 0.1s linear; -otransition: background-position 0.1s linear; transition: background-position 0.1s linear; }

.btn-primary, .btn-primary:hover { text-shadow: 0 -1px 0 rgba(0, 0, 0, 0.25); color: #ffffff; }

.btn-primary.active { color: rgba(255, 255, 255, 0.75); }

.btn-primary { background-color: #4a77d4; background-image: -moz-linear-gradient(top, #6eb6de, #4a77d4); background-image: -ms-linear-gradient(top, #6eb6de, #4a77d4); background-image: -webkit-gradient(linear, 0 0, 0 100%, from(#6eb6de), to(#4a77d4)); background-image: -webkit-linear-gradient(top, #6eb6de, #4a77d4); background-image: -

```
olinear-gradient(top, #6eb6de, #4a77d4); background-image: linear-gradient(top, #6eb6de, #4a77d4); background-repeat: repeat-x; filter: progid:dximagetransform.microsoft.gradient(startColorstr=#6eb6de, endColorstr=#4a77d4, GradientType=0);  border: 1px solid #3762bc; text-shadow: 1px 1px 1px rgba(0,0,0,0.4); box-shadow: inset 0 1px 0 rgba(255, 255, 255, 0.2), 0 1px 2px rgba(0, 0, 0, 0.5); }

.btn-primary:hover, .btn-primary:active, .btn-primary.active, .btn-primary.disabled, .btnprimary[disabled] { filter: none; background-color: #4a77d4; }

.btn-block { width: 100%; display:block; }


* { -webkit-box-sizing:border-box; -moz-box-sizing:border-box; -ms-box-sizing:border-box; -o-box-sizing:border-box; box-sizing:border-box; }


html { width: 100%; height:100%; overflow:hidden; }


body {
        width: 100%;
        height:100%;
        font-family: 'Open Sans', sans-serif;
        background: #092756;          color:
#fff;    font-size: 18px;          text-
align:center;    letter-spacing:1.2px;
        background: -moz-radial-gradient(0% 100%, ellipse cover, rgba(104,128,138,.4) 10%,rgba(138,114,76,0) 40%),-moz-linear-gradient(top,  rgba(57,173,219,.25) 0%, rgba(42,60,87,.4) 100%), -moz-linear-gradient(-45deg,  #670d10 0%, #092756 100%);
        background: -webkit-radial-gradient(0% 100%, ellipse cover, rgba(104,128,138,.4) 10%,rgba(138,114,76,0) 40%), -webkit-linear-gradient(top,  rgba(57,173,219,.25) 0%,rgba(42,60,87,.4) 100%), -webkit-linear-gradient(-45deg,  #670d10 0%,#092756 100%);
        background: -o-radial-gradient(0% 100%, ellipse cover, rgba(104,128,138,.4) 10%,rgba(138,114,76,0) 40%), -o-linear-gradient(top,  rgba(57,173,219,.25) 0%,rgba(42,60,87,.4) 100%), -o-linear-gradient(-45deg,  #670d10 0%,#092756 100%);
        background: -ms-radial-gradient(0% 100%, ellipse cover, rgba(104,128,138,.4) 10%,rgba(138,114,76,0) 40%), -ms-linear-gradient(top,  rgba(57,173,219,.25) 0%,rgba(42,60,87,.4) 100%), -ms-linear-gradient(-45deg,  #670d10 0%,#092756 100%);
        background: -webkit-radial-gradient(0% 100%, ellipse cover, rgba(104,128,138,.4) 10%,rgba(138,114,76,0) 40%), linear-gradient(to bottom,  rgba(57,173,219,.25) 0%,rgba(42,60,87,.4) 100%), linear-gradient(135deg,  #670d10 0%,#092756 100%);
```

```
        filter: progid:DXImageTransform.Microsoft.gradient( startColorstr='#3E1D6D',
endColorstr='#092756',GradientType=1 );


}
.login {
        position: absolute;     top:
40%;   left: 50%;       margin: -
150px 0 0 -150px;
        width:400px;
        height:400px;
}


.login h1 { color: #fff; text-shadow: 0 0 10px rgba(0,0,0,0.3); letter-spacing:1px;
textalign:center; }


input {
        width: 100%;
        margin-bottom: 10px;
        background: rgba(0,0,0,0.3);
        border: none;   outline:
none;   padding: 10px;          font-
size: 13px;     color: #fff;
        text-shadow: 1px 1px 1px rgba(0,0,0,0.3);
        border: 1px solid rgba(0,0,0,0.3);      border-
radius: 4px;

        box-shadow: inset 0 -5px 45px rgba(100,100,100,0.2), 0 1px 1px
rgba(255,255,255,0.2);
        -webkit-transition: box-shadow .5s ease;
        -moz-transition: box-shadow .5s ease;
```

```
        -o-transition: box-shadow .5s ease;    -ms-
transition: box-shadow .5s ease;        transition: box-
shadow .5s ease;
}
input:focus { box-shadow: inset 0 -5px 45px rgba(100,100,100,0.4), 0 1px 1px
rgba(255,255,255,0.2); }
```