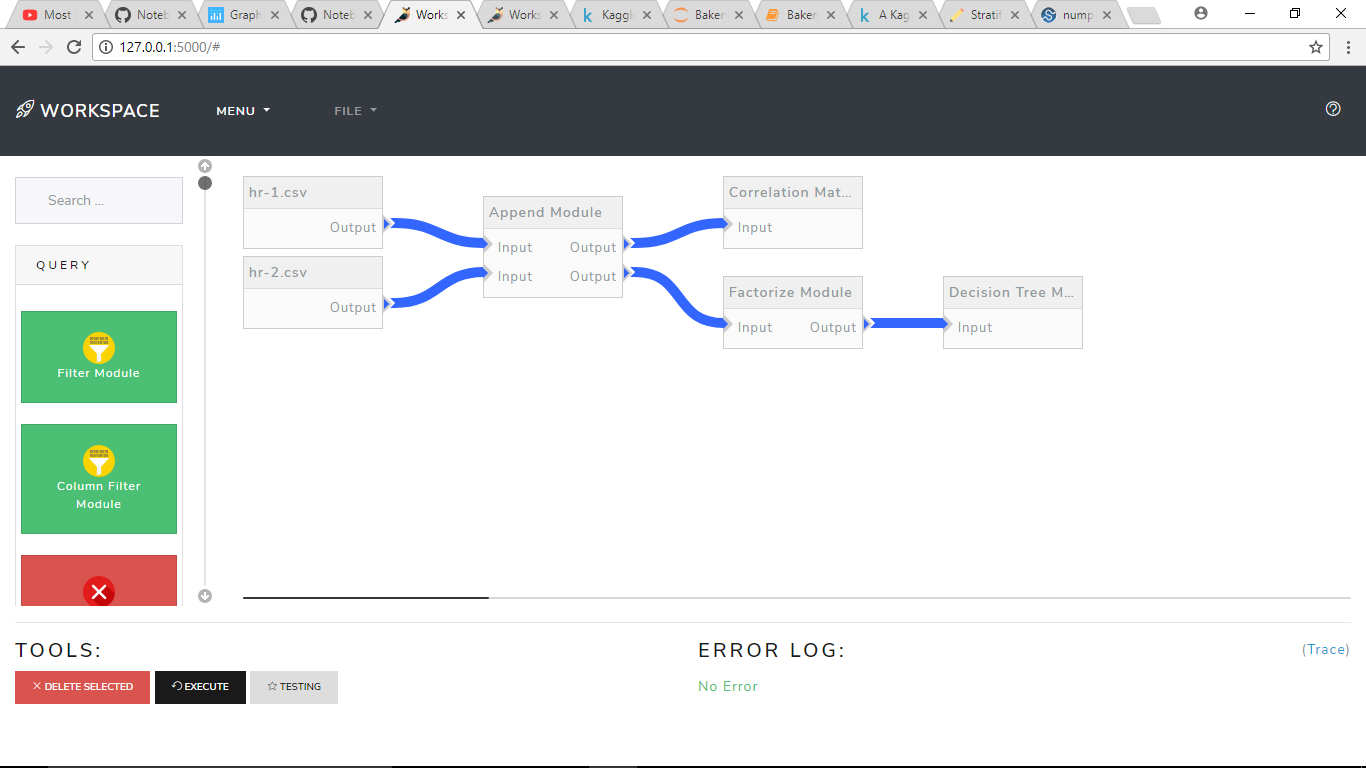
# PENGUJIAN SISTEM

Pada bab ini akan dijelaskan tentang pengujian terhadap aplikasi yang telah dibuat. Pengujian dilakukan dengan menggunakan 4 data dalam bentuk file dengan format CSV. Pengujian dilakukan pada *Directed Acyclic Graph* dengan parameternya, kamudian pengujian terhadap system rekomendasi metode analaisis.

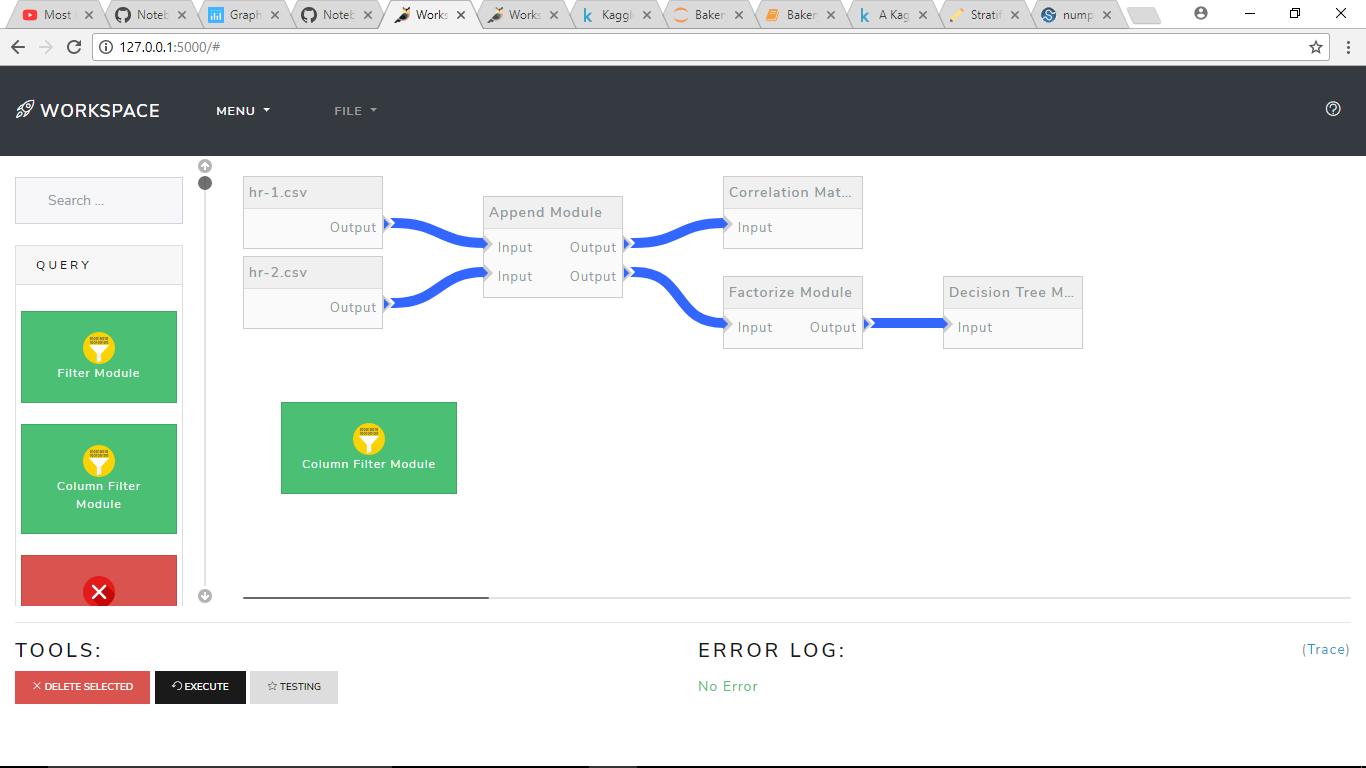
## Pengujian pada *Directed Acyclic Graph*

### Halaman *Workspace*



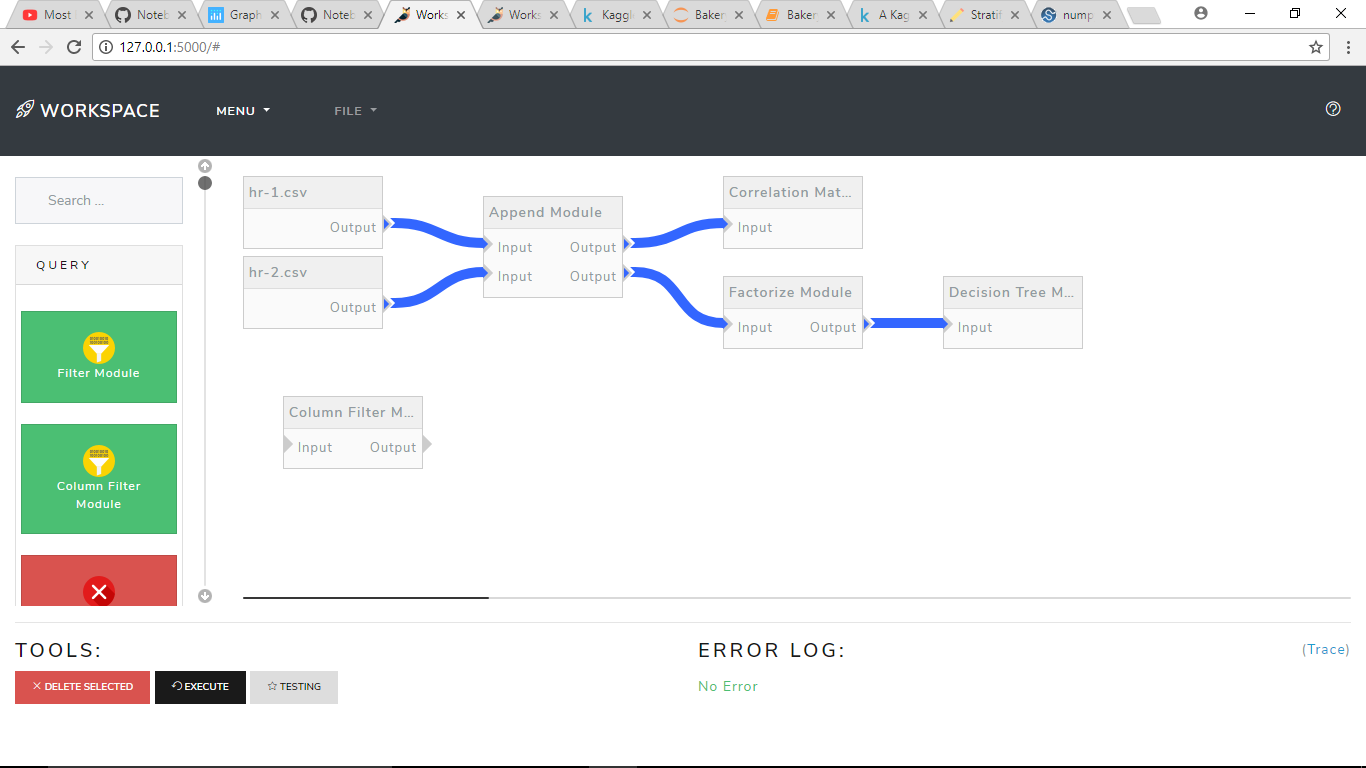
Halaman *Workspace*

Halaman ini merupakan halaman dimana *user* membuat / melakukan analisa sesusai dengan *data flow* yang diinginkan. Pada bagian kiri *user* dapat memilih *module* yang ingin ditambahkan kedalam *schema* yang ada. *Schema* merupakan *Directed Acyclic Graph* seperti Gambar 5.1.



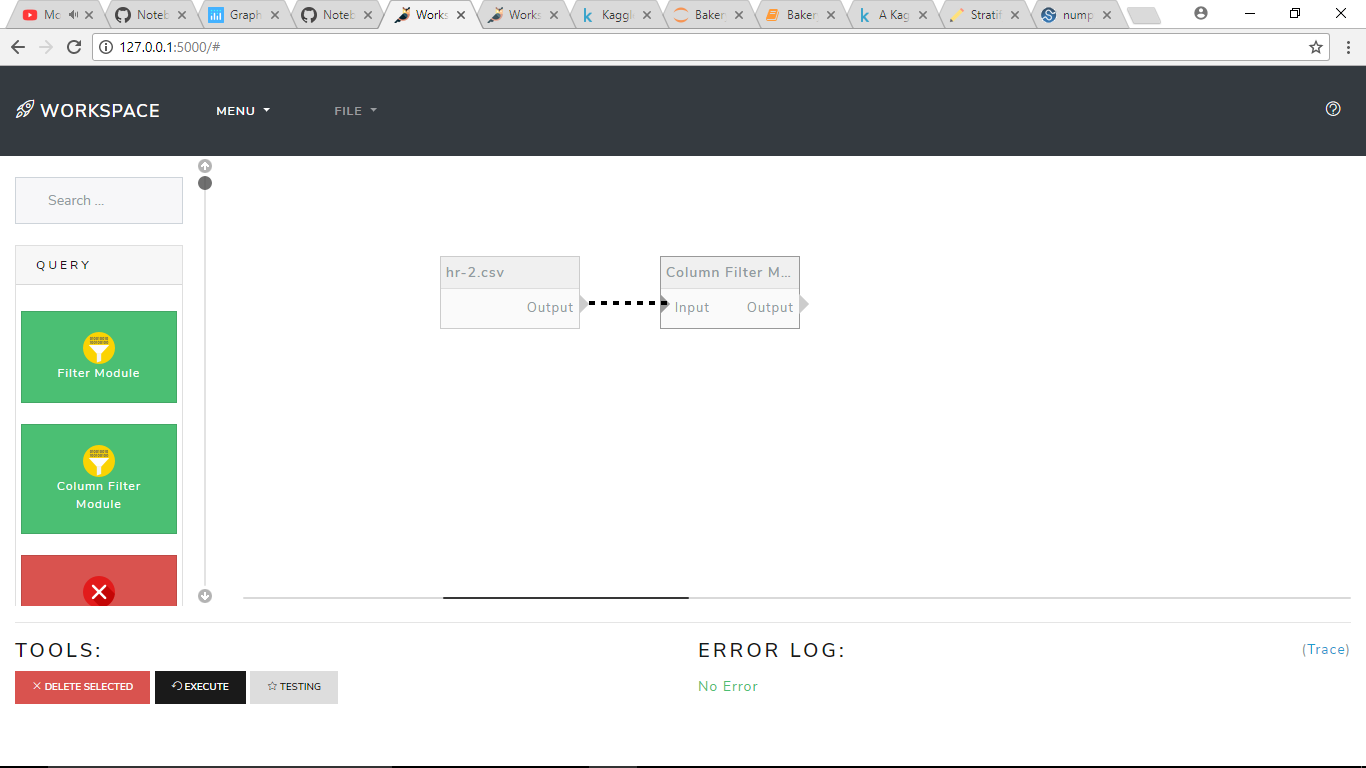
Tampilan *Drag* pada Halaman *Login*

*User* dapat memilih module dengan melakukan *drag* module menuju area *schema*. Setelah module ada pada posisi yang diingikan maka User dapat melakukan *drop*.



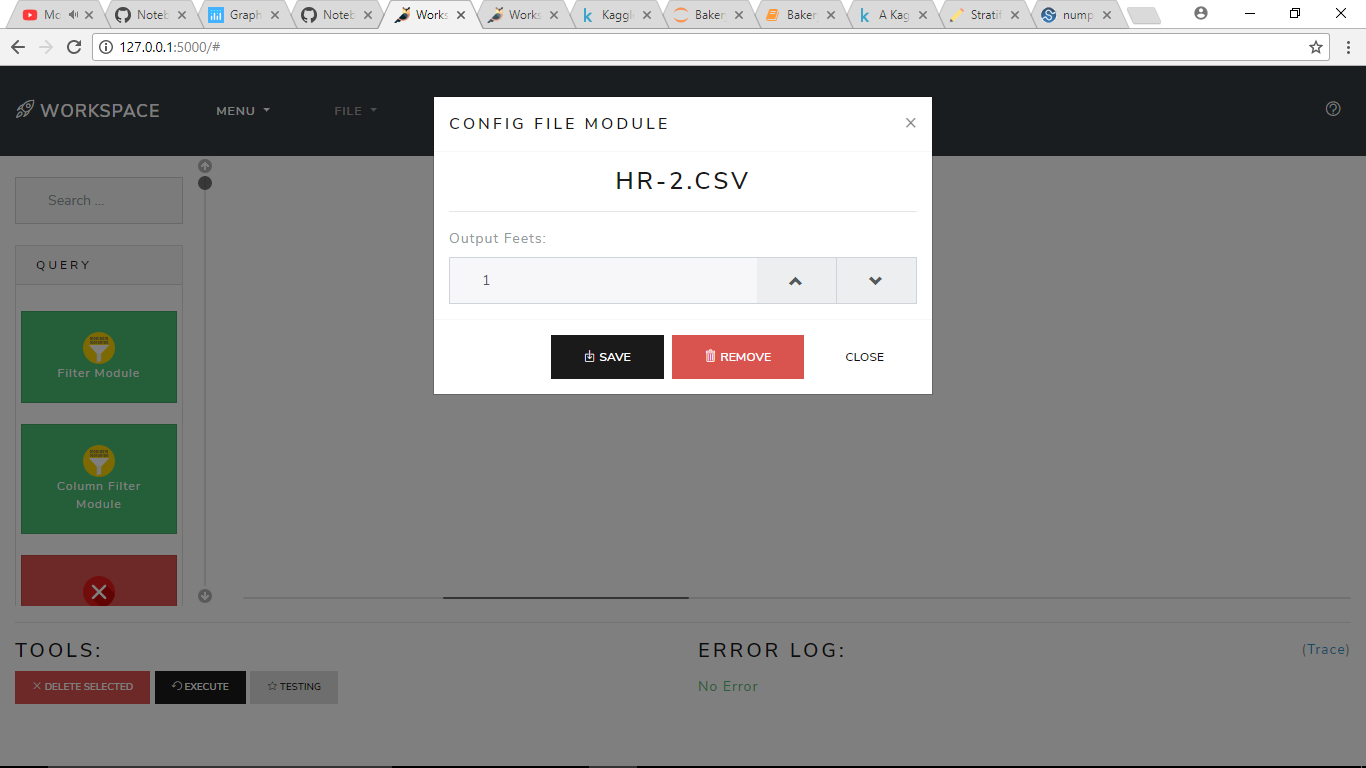
Tampilan *Drop* pada Halaman *Login*

Apabila *user* melakukan *drop* maka module akan terbuat sesuai dengan posisi *module* tersebut *drop*.



Fungsi *Data Flow*

Apabila *user* menekan tombol panah maka *user* dapat membuat sebuah *flow* atau alur data dari sebuah module menuju module lainya. *User* hanya perlu memilih panah *output* menuju panah *input* maka sebuah *flow* terbuat.

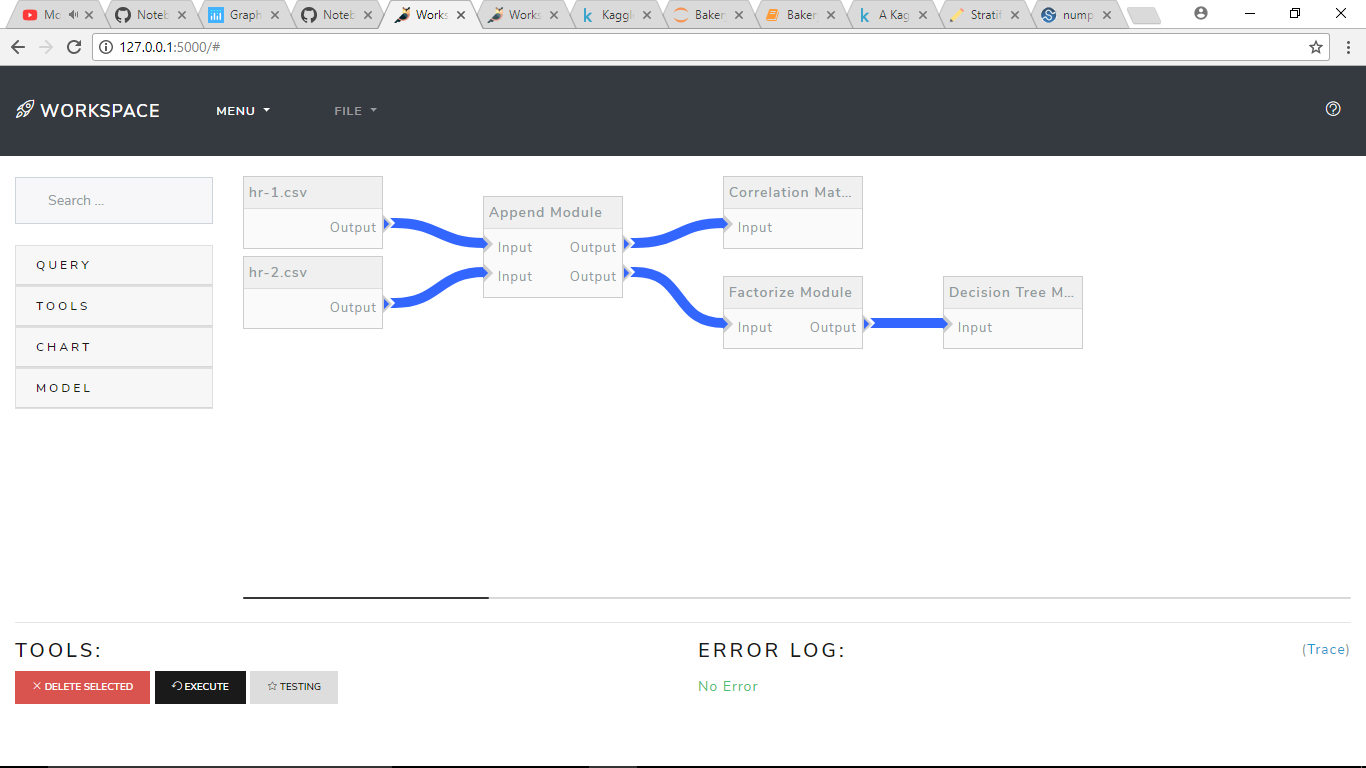


*Modal* untuk Parameter *Module*

Apabila *user* menekan salah satu *module*, maka sebuah modal akan muncul dengan *form* yang berisi parameter apa saja yang dapat dirubah dari module tersebut. Tiap *module* memiliki bentuk *modal* yang berbeda-beda.

### Fungsi *Execute* Workspace

Fungsi ini digunakan ketika *user* ingin menjalankan *Directect Acyclic Graph* (DAG)yang telah terbuat. Ketika tombol ini ditekan maka aplikasi akan mengirimkan *schema* berupa JSON untuk bentuk DAG yang ada, dan *parameters* berupa JSON yang mendukung data *schema* dengan menyediakan data parameter sesuai dengan *form* yang telah di isi *user* sebelumnya.



Tampilan *Directed Acyclic Graph* yang diuji

Pada Gambar 5.6 terdapat DAG yang digunakan untuk menguji apakah data yang di hasilkan sesuai data aslinya. Data yang digunakan:

1. hr-1.csv

Row: 1000

Data:

1. satisfaction\_level
2. last\_evaluation
3. number\_project
4. average\_montly\_hours
5. time\_spend\_company
6. Work\_accident
7. left
8. promotion\_last\_5years
9. department\_id
10. hr-2.csv

Row: 5000

Data:

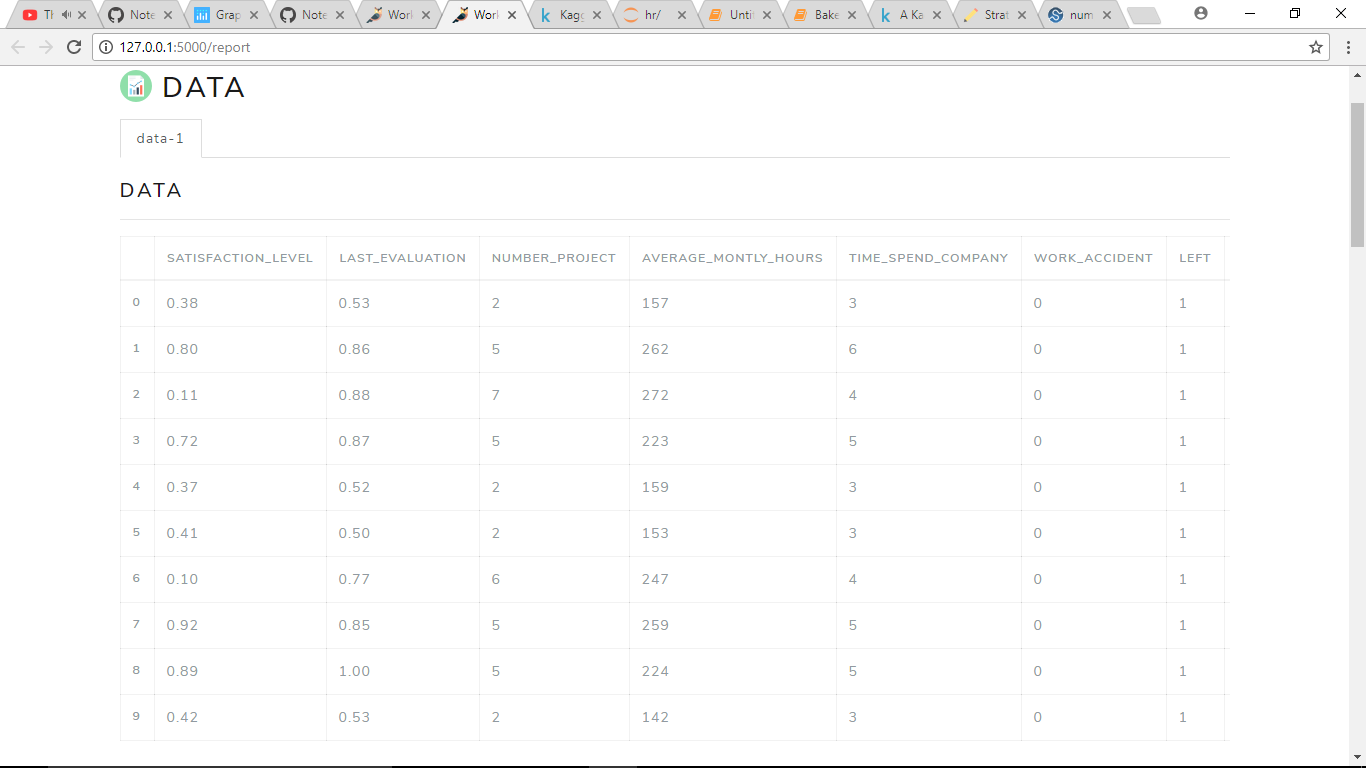
1. satisfaction\_level
2. last\_evaluation
3. number\_project
4. average\_montly\_hours
5. time\_spend\_company
6. Work\_accident
7. left
8. promotion\_last\_5years
9. department\_id

Setelah DAG terbentuk selanjutnya eksekusi akan dilakukan untuk melihat hasilnya. Pada DAG terdapat module *Correlation Matrix*, module ini akan memberikan hasil *image* *Correlation Matrix* yang dapat di-*download.* Pada DAG juga terdapat *module* *Decision Tree,* maka aplikasi akan memberikan statistic sederhana mengenai *Decision Tree* yang terbuat.

Hasil yang didapatkan jika menekan tombol *execute* adalah sebuah halaman yang terbagi menjadi 3 bagian.

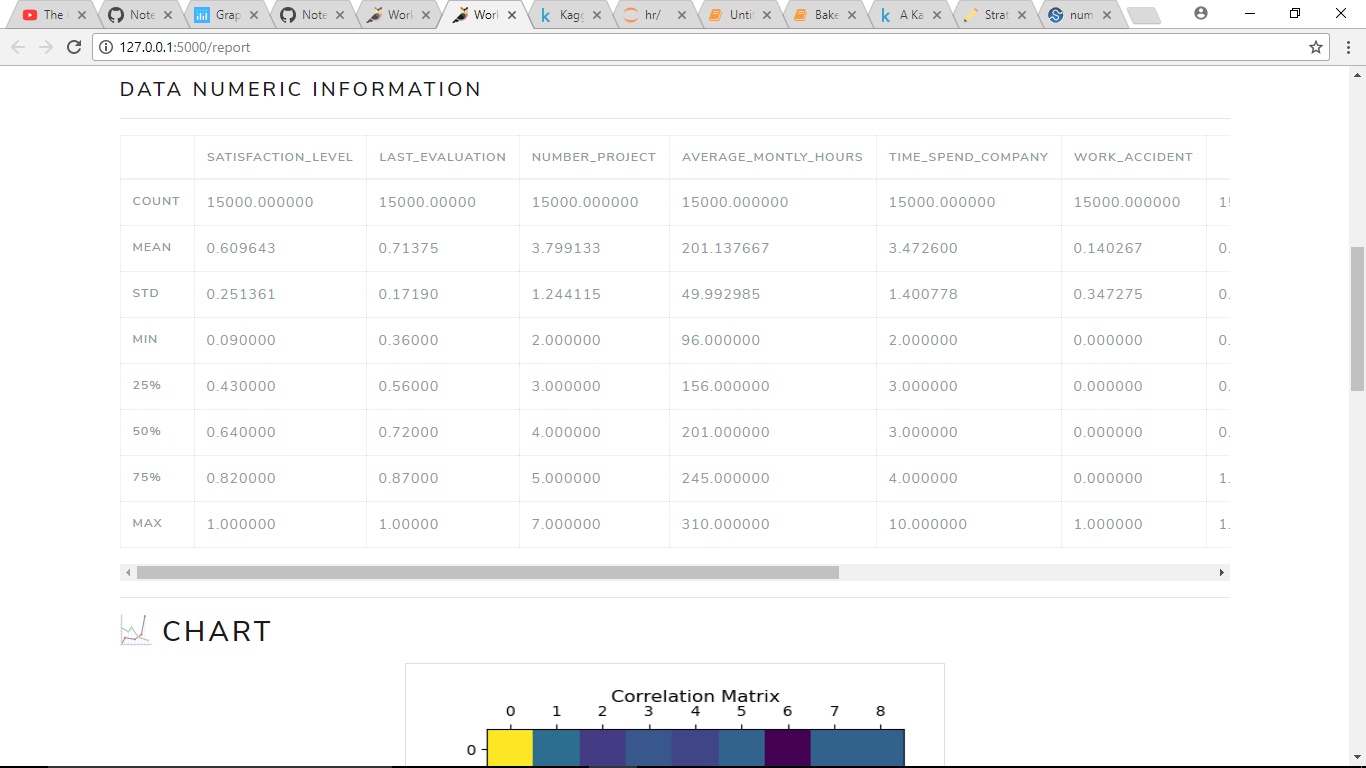
1. Data

Merupakan bagian yang menampilkan data yang dihasilkan pada akhir *data flow* yang berjalan.



Hasil *Execute* Bagian Data

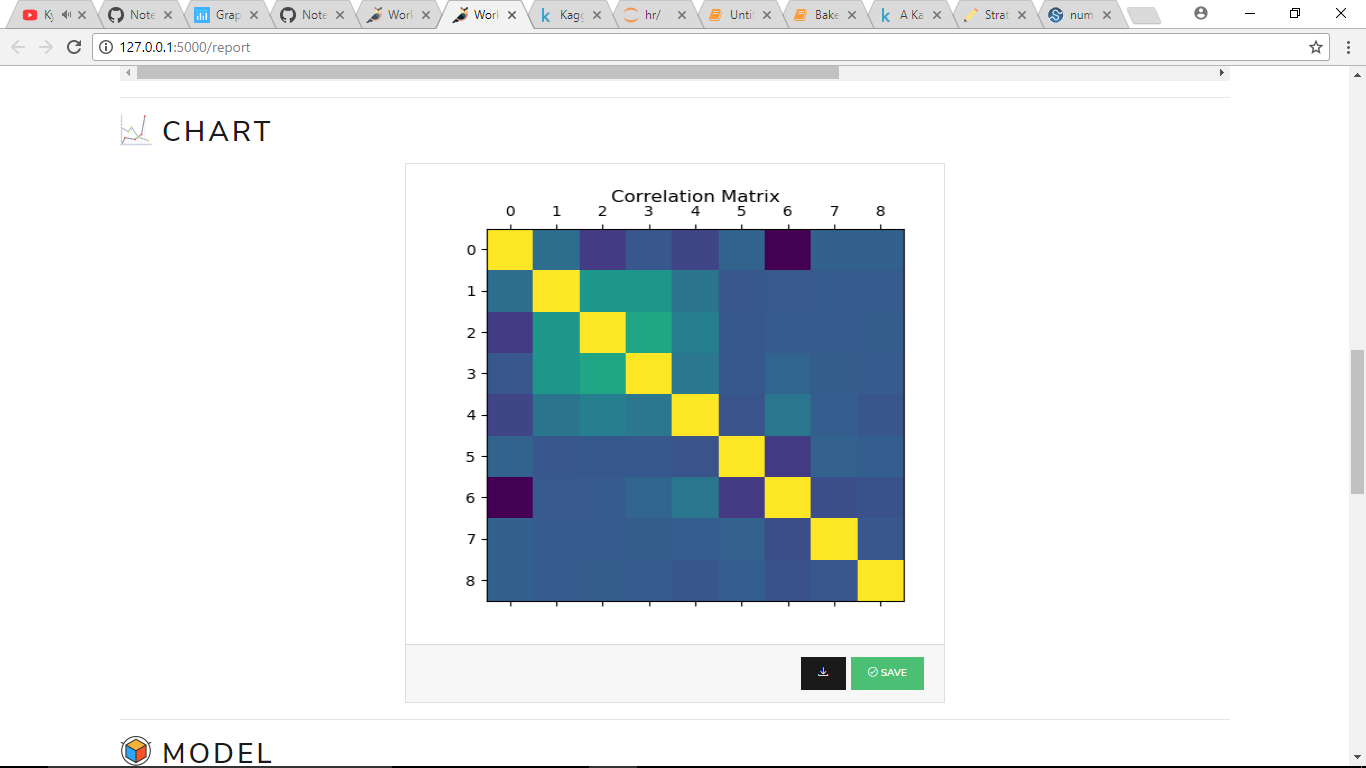
Selain itu bagian ini juga memberikan deskripsi singkat mengenai data yang ada.



Hasil *Execute* *Data Description*

1. Chart

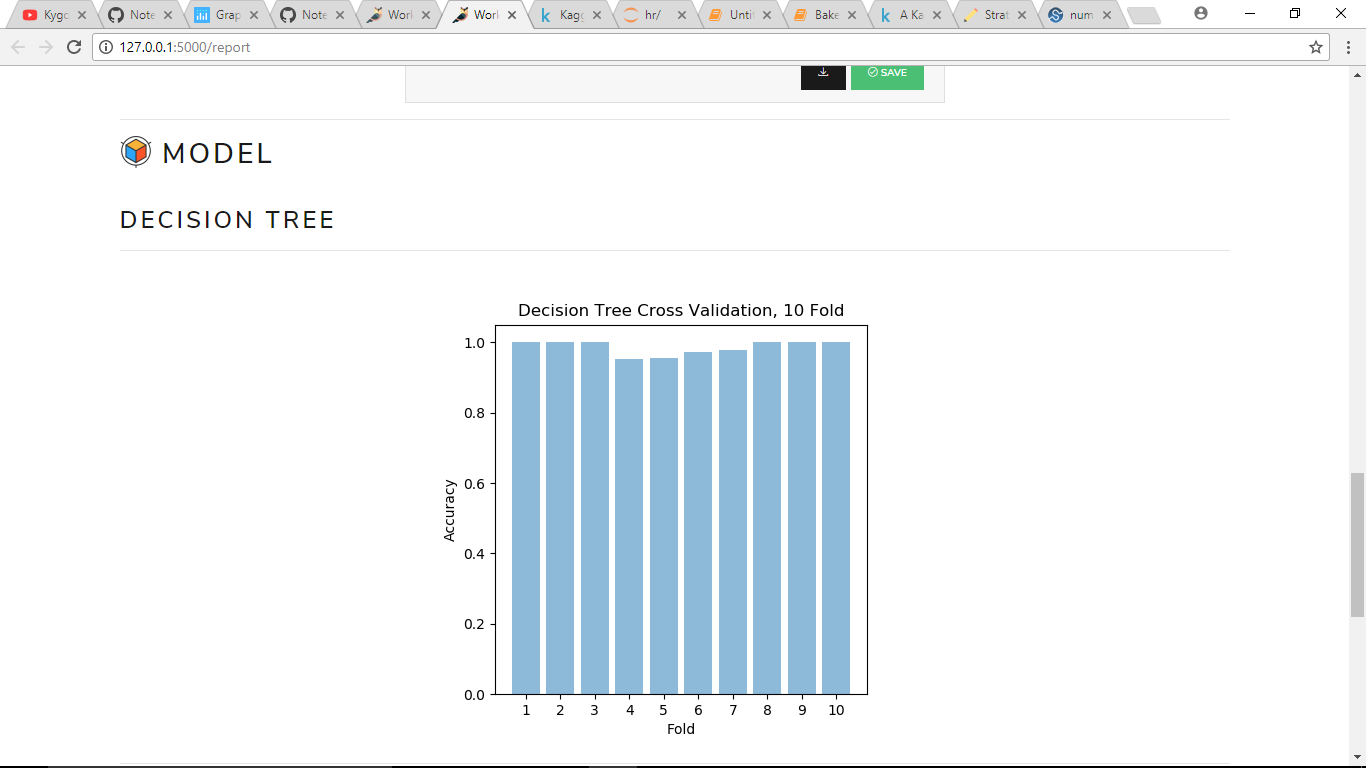
Merupakan bagian yang memberikan gambar *chart – chart*  yang tergenerasi dari aplikasi melalui DAG yang ada. Gambar yang dihasilkan dapat di-*download.*



Hasil *Execute* Bagian *Chart*

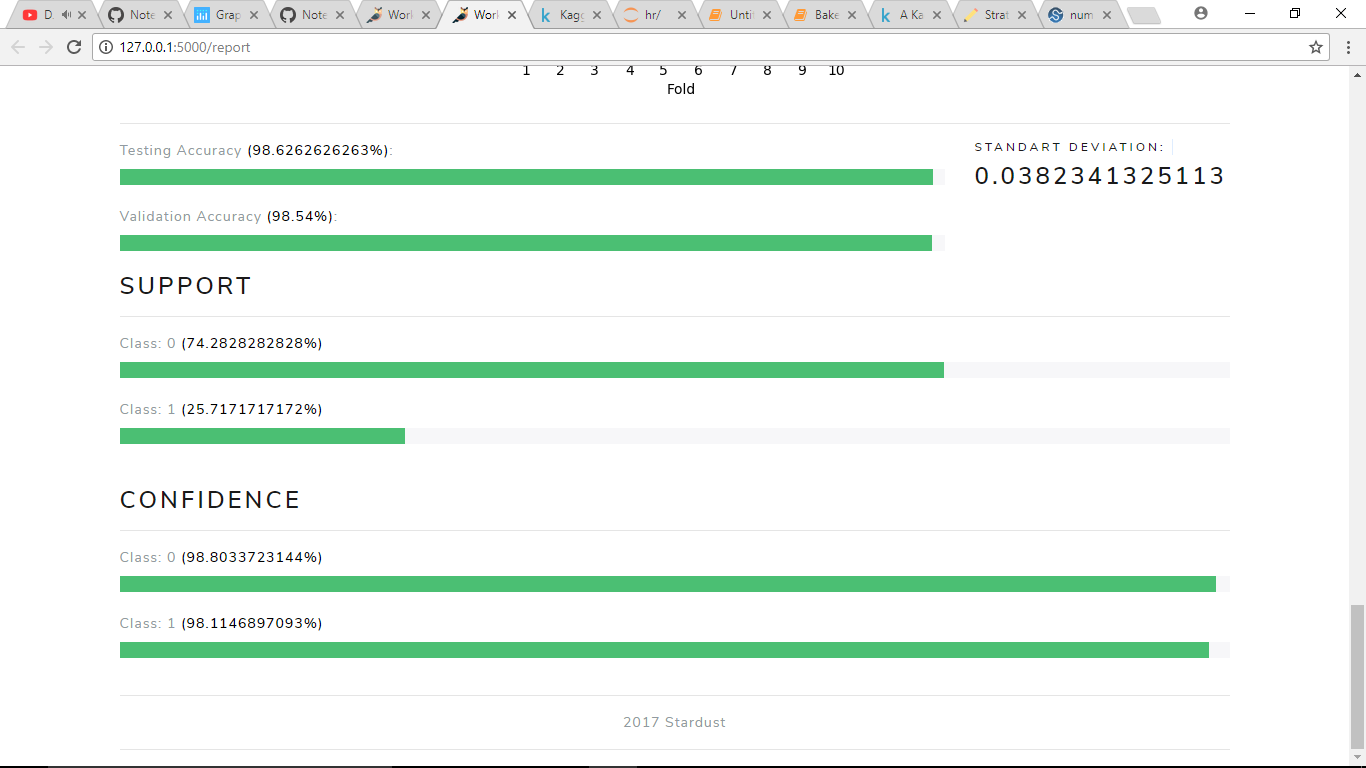
1. Model

Merupakan bagian yang menampilkan statistik singkat mengenai model yang tergenerasi. Pada bagian ini aplikasi akan memberikan sebuah grafik yang menunjukan hasil dari model yang sudah di-validasi menggunakan metode *cross validation.*



Hasil *Execute* Bagian Model

Selanjutnya aplikasi akan memberikan *accuracy* rata – rata dari *cross validation* yang sudah dilakukan. Selain itu terdapat juga hasil *accuracy testing* yang didapat dengan membagi data menjadi 70% *training* dan 30% *testing.* Selain itu terdapat *standart deviation* yang digunakan untuk membantu pengguna seberapa konsisten model yang terbuat. Aplikasi juga memberikan data *support* yang memberikan persentase *class* terhadap total data yang ada dan *confidence* yang memberikan persentase seberapa banyak data dari masing - masing *class* yang dapat diprediksi oleh model yang dibuat.



Statisik Singkat *Model*

## Hasil Pengujian *Recommender & Analyze*

*Recommender & Analyze* merupakan *module* yang ada dalam aplikasi dimana memiliki tugas untuk memberikan *user* sebuah rekomendasi metode analisa yang digunakan. Fitur ini dapat membantu *user* yang baru dalam melakukan *business intelligence*. Fitur ini akan mengambil *data sample* kemudian akan menjalankan 3 *Tread* untuk melakukan perhitungan data sesuai dengan 3 metode yang berbeda. Pengujian ini akan dilakukan dengan 4 data yang berbeda.

### Human Resource

Type: *Classification*

Sampling: *Random Sampling & Stratified*

Data Size: 14999 Rows

All Data Result (menjalankan model dengan seluruh data):

1. Decision Tree
2. Naïve Bayes
3. Logistic Regression

Tabel Iterasi *Random* *Sampling Human Resource*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Acc. Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 0.945288591 | 0.722501119 | 0.779176734 | 0.333333333 | 0.173430242 |
| 0.2 | 0.963971966 | 0.68192784 | 0.763262607 | 0.333333333 | 0.268932319 |
| 0.3 | 0.969327889 | 0.792620144 | 0.761504083 | 1 | 0.366500965 |
| 0.4 | 0.970160789 | 0.785457197 | 0.764461054 | 1 | 0.46339114 |
| 0.5 | 0.970526548 | 0.811177609 | 0.76637722 | 1 | 0.575948467 |
| **Average** | 0.963855157 | 0.758736781 | 0.766956339 | 0.733333333 |  |
| **Deviation** | 0.009580713 | 0.048630731 | 0.006313012 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 0.955297539 | 0.758899329 | 0.777190157 | 0.333333333 | 0.178435755 |
| 0.2 | 0.963315341 | 0.795926625 | 0.76126757 | 1 | 0.266813333 |
| 0.3 | 0.96132633 | 0.748850732 | 0.777498207 | 0.333333333 | 0.372229667 |
| 0.4 | 0.9734958 | 0.806133067 | 0.773623573 | 1 | 0.441946535 |
| 0.5 | 0.971464174 | 0.797307165 | 0.761966889 | 1 | 0.552007337 |
| **Average** | 0.964979837 | 0.781423384 | 0.770309279 | 0.733333333 |  |
| **Deviation** | 0.006699758 | 0.022984854 | 0.007229933 |  |  |
| Iteration-3 | | | | | |
| 0.1 | 0.9519277 | 0.813791102 | 0.787817355 | 1 | 0.173624144 |
| 0.2 | 0.962649804 | 0.745288625 | 0.772274385 | 0.333333333 | 0.276134158 |
| 0.3 | 0.966885385 | 0.813742443 | 0.764384323 | 1 | 0.354734564 |
| 0.4 | 0.971995764 | 0.80845813 | 0.772128322 | 1 | 0.489753654 |
| 0.5 | 0.975596603 | 0.774903969 | 0.764499524 | 1 | 0.563562975 |
| **Average** | 0.965811051 | 0.791236854 | 0.772220782 | 0.866666667 |  |
| **Deviation** | 0.00822074 | 0.027165312 | 0.008535718 |  |  |
| Iteration-4 | | | | | |
| 0.1 | 0.951964206 | 0.84255481 | 0.778519016 | 1 | 0.174648265 |
| 0.2 | 0.963010741 | 0.694372482 | 0.758876476 | 0.333333333 | 0.271207452 |
| 0.3 | 0.964218626 | 0.792824952 | 0.760823351 | 1 | 0.363709006 |
| 0.4 | 0.97416305 | 0.79313133 | 0.765962437 | 1 | 0.460978919 |
| 0.5 | 0.972396955 | 0.786509497 | 0.760627519 | 1 | 0.560321634 |
| **Average** | 0.965150716 | 0.781878614 | 0.76496176 | 0.866666667 |  |
| **Deviation** | 0.007912806 | 0.048180525 | 0.007179739 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 0.956038194 | 0.74734963 | 0.765181623 | 0.333333333 | 0.174790117 |
| 0.2 | 0.964324348 | 0.741232495 | 0.776261458 | 0.333333333 | 0.279952071 |
| 0.3 | 0.966873927 | 0.754621131 | 0.755497176 | 0.333333333 | 0.369000056 |
| 0.4 | 0.970661102 | 0.767627713 | 0.771297997 | 0.333333333 | 0.468762462 |
| 0.5 | 0.969062361 | 0.789436893 | 0.772901584 | 1 | 0.570975576 |
| **Average** | 0.965391986 | 0.760053572 | 0.768227968 | 0.466666667 |  |
| **Deviation** | 0.005139351 | 0.017122393 | 0.007309692 |  |  |

Tabel Iterasi *Stratified* *Sampling Human Resource*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 0.972898198 | 0.787099654 | 0.76200818 | 1 | 0.376808157 |
| 0.2 | 0.966892213 | 0.802890573 | 0.7602361 | 1 | 0.384331088 |
| 0.3 | 0.967556937 | 0.800876911 | 0.756001253 | 1 | 0.382396474 |
| 0.4 | 0.957113456 | 0.797348559 | 0.769789451 | 1 | 0.383616329 |
| 0.5 | 0.971998465 | 0.776001033 | 0.769558053 | 1 | 0.383825992 |
| **Average** | 0.967291854 | 0.792843346 | 0.763518607 | 1 |  |
| **Deviation** | 0.005610936 | 0.010024324 | 0.005391943 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 0.970896153 | 0.798246905 | 0.762459568 | 1 | 0.389196581 |
| 0.2 | 0.96511246 | 0.798651568 | 0.761767439 | 1 | 0.382742857 |
| 0.3 | 0.965555496 | 0.816876675 | 0.770665251 | 1 | 0.398653394 |
| 0.4 | 0.966446389 | 0.801105003 | 0.767776065 | 1 | 0.392024462 |
| 0.5 | 0.972439944 | 0.764675106 | 0.76600003 | 0.333333333 | 0.383468612 |
| **Average** | 0.968090088 | 0.795911051 | 0.76573367 | 0.866666667 |  |
| **Deviation** | 0.002992906 | 0.017059809 | 0.003317033 |  |  |
| Iteration-3 | | | | | |
| 0.1 | 0.964888889 | 0.793777778 | 0.760444444 | 1 | 0.397228274 |
| 0.2 | 0.969109539 | 0.803110041 | 0.769113392 | 1 | 0.396800518 |
| 0.3 | 0.964663099 | 0.780197842 | 0.774001788 | 1 | 0.388273259 |
| 0.4 | 0.967778743 | 0.806209185 | 0.778438292 | 1 | 0.412225386 |
| 0.5 | 0.964221195 | 0.784893119 | 0.751774063 | 1 | 0.38406791 |
| **Average** | 0.966132293 | 0.793637593 | 0.766754396 | 1 |  |
| **Deviation** | 0.00194585 | 0.010048727 | 0.009579445 |  |  |
| Iteration-4 | | | | | |
| 0.1 | 0.963331787 | 0.818898898 | 0.776445983 | 1 | 0.392411531 |
| 0.2 | 0.9662216 | 0.782885545 | 0.763790663 | 1 | 0.402125657 |
| 0.3 | 0.971332262 | 0.808207327 | 0.767098071 | 1 | 0.435058456 |
| 0.4 | 0.96822651 | 0.796862207 | 0.771768256 | 1 | 0.408207707 |
| 0.5 | 0.966442927 | 0.786899032 | 0.761106908 | 1 | 0.369283394 |
| **Average** | 0.967111017 | 0.798750602 | 0.768041976 | 1 |  |
| **Deviation** | 0.002629951 | 0.013349976 | 0.005505856 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 0.966890341 | 0.782229908 | 0.767553341 | 1 | 0.39026982 |
| 0.2 | 0.969775279 | 0.805108954 | 0.766884222 | 1 | 0.382185345 |
| 0.3 | 0.964894641 | 0.70514768 | 0.753559771 | 0.333333333 | 0.395452373 |
| 0.4 | 0.967990029 | 0.78799197 | 0.777337099 | 1 | 0.387301554 |
| 0.5 | 0.964667518 | 0.788209037 | 0.765119094 | 1 | 0.373672014 |
| **Average** | 0.966843562 | 0.77373751 | 0.766090705 | 0.866666667 |  |
| **Deviation** | 0.001920677 | 0.035138207 | 0.007579165 |  |  |

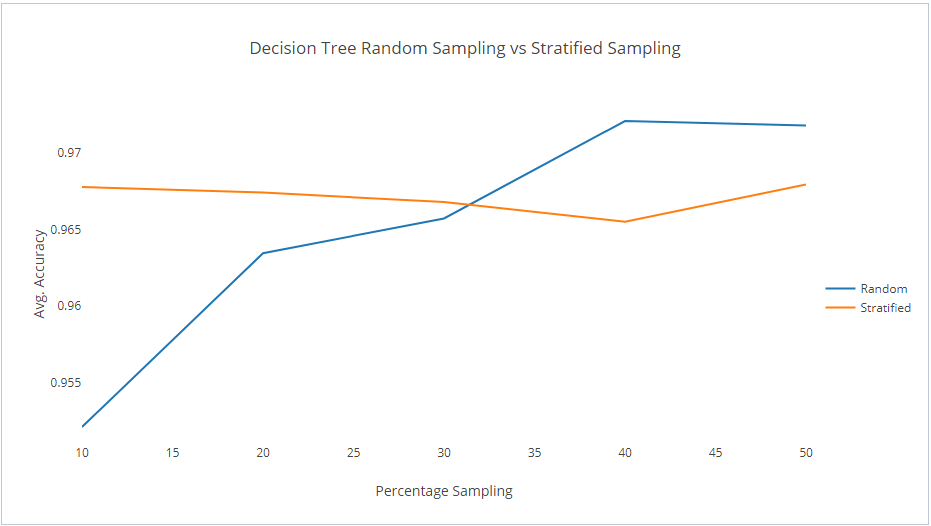
Tabel Rata – Rata Hasil *Random Sampling* *Human Resource*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Acc. Recommender | Time |
| 0.1 | 0.952103246 | 0.777019198 | 0.777576977 | 0.6 | 0.174985705 |
| 0.2 | 0.96345444 | 0.731749614 | 0.766388499 | 0.466666667 | 0.272607866 |
| 0.3 | 0.965726431 | 0.78053188 | 0.763941428 | 0.733333333 | 0.365234852 |
| 0.4 | 0.972095301 | 0.792161487 | 0.769494677 | 0.866666667 | 0.464966542 |
| 0.5 | 0.971809328 | 0.791867026 | 0.765274547 | 1 | 0.564563198 |
| **Average** | 0.965037749 | 0.774665841 | 0.768535226 |  |  |
| **Deviation** | 0.007510674 | 0.032816763 | 0.007313619 |  |  |

Tabel Rata – Rata Hasil *Stratified Sampling* *Human Resource*

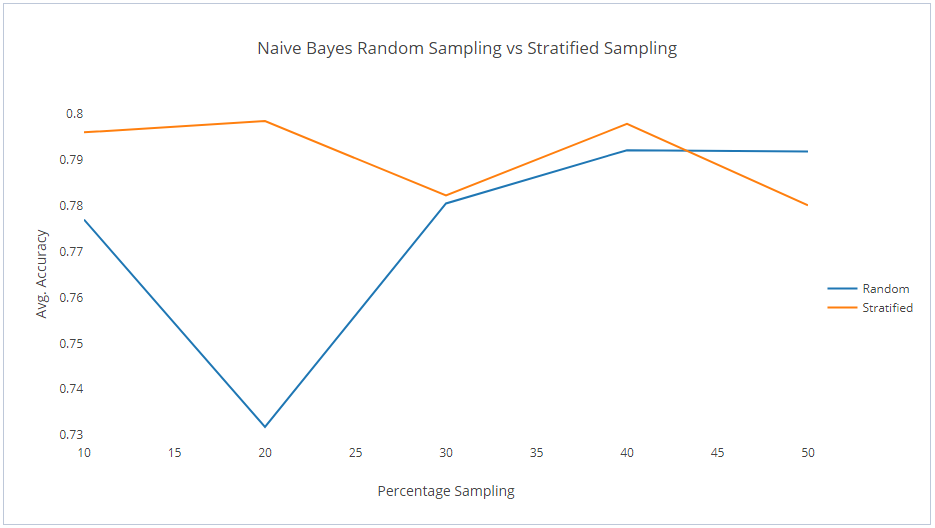
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Recommender | Time |
| 0.1 | 0.967781073 | 0.796050629 | 0.765782303 | 1 | 0.389182873 |
| 0.2 | 0.967422218 | 0.798529336 | 0.764358363 | 1 | 0.389637093 |
| 0.3 | 0.966800487 | 0.782261287 | 0.764265227 | 0.866666667 | 0.399966791 |
| 0.4 | 0.965511026 | 0.797903385 | 0.773021832 | 1 | 0.396675087 |
| 0.5 | 0.96795401 | 0.780135465 | 0.76271163 | 0.866666667 | 0.378863584 |
| **Average** | 0.967093763 | 0.79097602 | 0.766027871 | 0.946666667 | 0.390865086 |
| **Deviation** | 0.003020064 | 0.017124209 | 0.006274688 | 0.065319726 | 0.007279712 |

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Decision Tree.* Perbandingan tersebut dapat dibentuk dalam bentuk *chart* pada Gambar 5.12



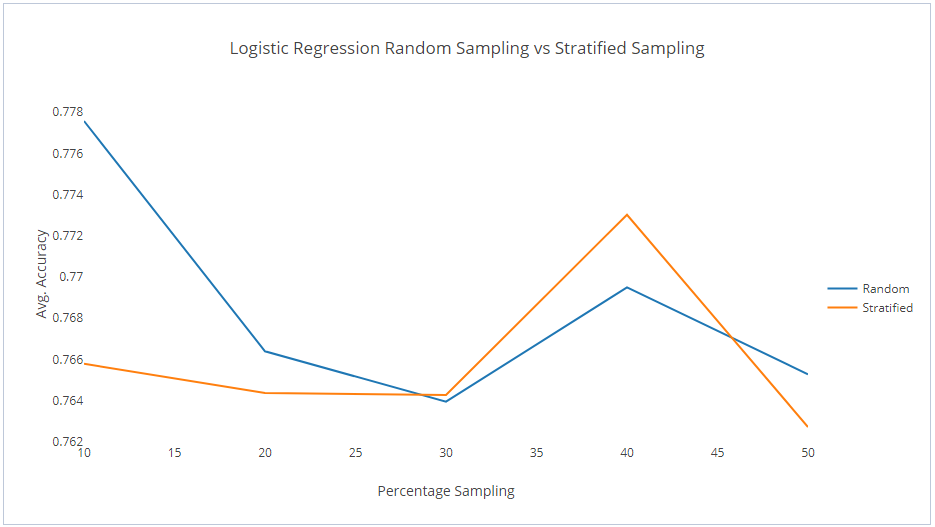
Grafik Perbandingan *Accuracy* *Decision Tree*

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Naïve Bayes.* Perbandingan tersebut dapat dibentuk dalam bentuk *chart* pada Gambar 5.13



Grafik Perbandingan *Accuracy* *Naïve Bayes*

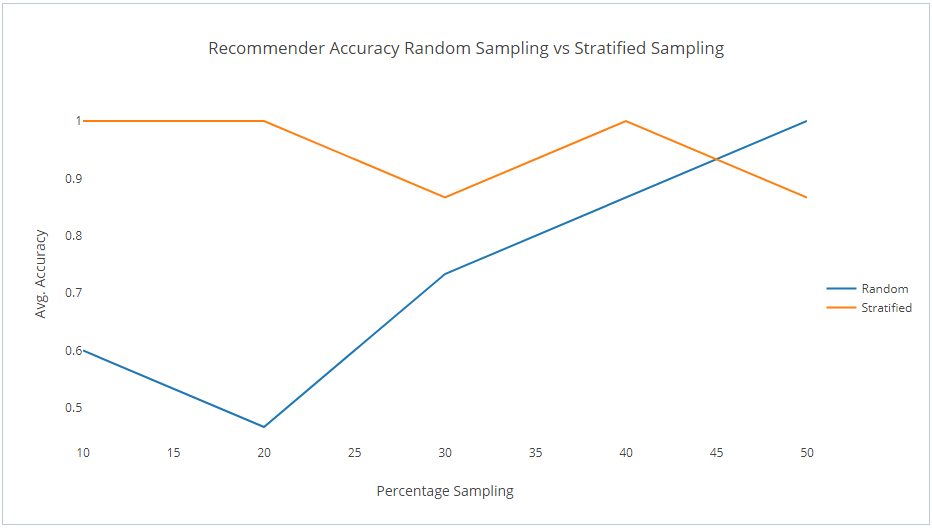
Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Logistic Regression.* Perbandingan tersebut dapat dibentuk dalam bentuk *chart* pada Gambar 5.14



Grafik Perbandingan *Accuracy* *Logistic Regression*

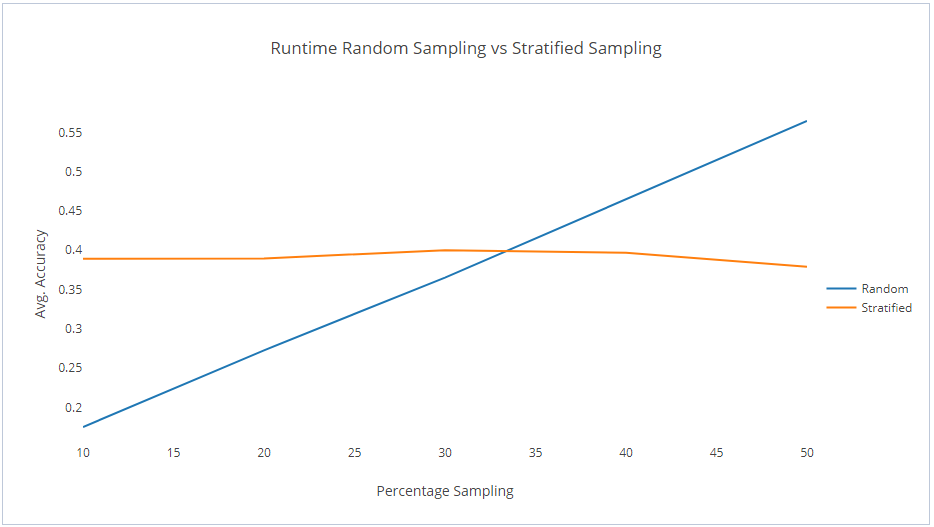
Beberapa analisa yang dapat dilihat dari grafik Gambar 5.12, Gambar 5.13, Gambar 5.14, system rekomendasi dengan menggunakan metode *Stratified Sampling* lebih konsisten, dimana grafik banyak menunjukan garis yang lurus dibandingkan dengan menggunakan *Random Sampling.* Kemudian semakin banyak data, cenderung metode prediksi memberikan prediksi lebih baik.

Berdasarkan Tabel 5.2 dan Tabel 5.3 hasil *accuracy* sistem rekomendasi pada ukuran data sample dari 10 % sampai 50%, dapat dibentuk grafik yang membandingkan *accuracy* sistem rekomendasi pada *Random Sampling*  dan *Stratified Sampling*.



Grafik Perbandingan *Accuracy* Sistem Rekomendasi

Selain itu waktu yang diperlukan sistem rekomendasi untuk menjalankan proses rekomendasi pada data *human resource* dapat dibentuk dengan Gambar 5.16.



*Runtime* Sistem Rekomendasi

Berdasarkan Gambar 5.15, sistem rekomendasi menggunakan *Stratified Sampling* lebih konsisten dibandingkan dengan menggunakan metode *Random Sampling.* Namun berdasarkan Gambar 5.16 kecepatan waktu proses sistem rekomendasi dengan menggunakan *Random Sampling*  memiliki waktu eksekusi yang lebih cepat dibandingkan dengan menggunakan *Stratified Sampling.* Tetapi ketika proses sudah mencapat data dengan ukuran yang lebih besar, waktu eksekusi *Random Sampling* menjadi lebih lama dibandingkan dengan *Stratified Sampling.*

### *West Nile*

Type: *Classification*

Sampling: *Random Sampling & Stratified*

Data Size: 10506 Rows

All Data Result (menjalankan model dengan seluruh data):

1. Logistic Regression
2. Naïve Bayes
3. Decision Tree

Tabel Iterasi *Random* *Sampling West Nile*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 0.920952381 | 0.923809524 | 0.951428571 | 1 | 0.2248808 |
| 0.2 | 0.908167025 | 0.859125938 | 0.938602765 | 0.333333333 | 0.322906526 |
| 0.3 | 0.907971119 | 0.881608597 | 0.95017753 | 0.333333333 | 0.523330836 |
| 0.4 | 0.921004171 | 0.889358338 | 0.951215764 | 0.333333333 | 0.650989373 |
| 0.5 | 0.909577404 | 0.876454101 | 0.946126743 | 0.333333333 | 0.759120692 |
| **Average** | 0.91353442 | 0.8860713 | 0.947510275 | 0.466666667 |  |
| **Deviation** | 0.006103123 | 0.0213234 | 0.004846769 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 0.919064206 | 0.92387138 | 0.951436692 | 1 | 0.218383457 |
| 0.2 | 0.915763612 | 0.870541142 | 0.946695768 | 0.333333333 | 0.32036235 |
| 0.3 | 0.909560827 | 0.901644424 | 0.952714397 | 0.333333333 | 0.45921768 |
| 0.4 | 0.911472387 | 0.884095267 | 0.948359744 | 0.333333333 | 0.520599723 |
| 0.5 | 0.913006585 | 0.873206031 | 0.948792251 | 0.333333333 | 0.764571922 |
| **Average** | 0.913773523 | 0.890671649 | 0.94959977 | 0.466666667 |  |
| **Deviation** | 0.003334127 | 0.019880432 | 0.00217707 |  |  |

Tabel Iterasi *Random* *Sampling West Nile* (Lanjutan)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Iteration-3 | | | | | |
| 0.1 | 0.909627479 | 0.911559196 | 0.949548172 | 1 | 0.224664906 |
| 0.2 | 0.911962727 | 0.84385777 | 0.941940374 | 0.333333333 | 0.384250082 |
| 0.3 | 0.902257006 | 0.859417761 | 0.940338327 | 0.333333333 | 0.474749944 |
| 0.4 | 0.920282595 | 0.874581193 | 0.950500967 | 0.333333333 | 0.678803031 |
| 0.5 | 0.916048036 | 0.885024275 | 0.949553794 | 0.333333333 | 0.763930105 |
| **Average** | 0.912035569 | 0.874888039 | 0.946376327 | 0.466666667 |  |
| **Deviation** | 0.006091598 | 0.023020939 | 0.00431983 |  |  |
| Iteration-4 | | | | | |
| 0.1 | 0.913313982 | 0.840957737 | 0.94191271 | 0.333333333 | 0.196304732 |
| 0.2 | 0.914341401 | 0.889143799 | 0.95003572 | 0.333333333 | 0.379682588 |
| 0.3 | 0.90479468 | 0.906388054 | 0.94668555 | 1 | 0.532041733 |
| 0.4 | 0.917900988 | 0.890763116 | 0.945979915 | 0.333333333 | 0.557720278 |
| 0.5 | 0.91624219 | 0.86768415 | 0.946507293 | 0.333333333 | 0.649446593 |
| **Average** | 0.913318648 | 0.878987371 | 0.946224238 | 0.466666667 |  |
| **Deviation** | 0.004544205 | 0.022656718 | 0.002587249 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 0.934293835 | 0.890465305 | 0.953341627 | 0.333333333 | 0.219408678 |
| 0.2 | 0.911930181 | 0.885793566 | 0.944797064 | 0.333333333 | 0.36195693 |
| 0.3 | 0.91622098 | 0.884151114 | 0.950495899 | 0.333333333 | 0.435238062 |
| 0.4 | 0.919324459 | 0.867466521 | 0.946693635 | 0.333333333 | 0.606414431 |
| 0.5 | 0.9128158 | 0.852270771 | 0.944603979 | 0.333333333 | 0.858878701 |
| **Average** | 0.918917051 | 0.876029455 | 0.947986441 | 0.333333333 |  |
| **Deviation** | 0.008123077 | 0.014195013 | 0.0034135 |  |  |

Tabel Iterasi *Stratified* *Sampling West Nile*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 0.908636741 | 0.92202571 | 0.948648835 | 1 | 0.20788273 |
| 0.2 | 0.908199052 | 0.886289777 | 0.947192507 | 0.333333333 | 0.37179678 |
| 0.3 | 0.911491344 | 0.868659061 | 0.948285803 | 0.333333333 | 0.51472367 |
| 0.4 | 0.912922746 | 0.880561588 | 0.947893903 | 0.333333333 | 0.673233042 |
| 0.5 | 0.919093849 | 0.876642957 | 0.94764939 | 0.333333333 | 0.788466134 |
| **Average** | 0.912068746 | 0.886835819 | 0.947934088 | 0.466666667 |  |
| **Deviation** | 0.003928561 | 0.018505575 | 0.000503424 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 0.920102288 | 0.92202571 | 0.948648835 | 1 | 0.208140043 |
| 0.2 | 0.910582261 | 0.886289777 | 0.947192507 | 0.333333333 | 0.371851028 |
| 0.3 | 0.908634176 | 0.868659061 | 0.948285803 | 0.333333333 | 0.512713548 |
| 0.4 | 0.911735098 | 0.880561588 | 0.947893903 | 0.333333333 | 0.667174451 |
| 0.5 | 0.919854303 | 0.876642957 | 0.94764939 | 0.333333333 | 0.801785212 |
| **Average** | 0.914181625 | 0.886835819 | 0.947934088 | 0.466666667 |  |
| **Deviation** | 0.004836292 | 0.018505575 | 0.000503424 |  |  |
| Iteration-3 | | | | | |
| 0.1 | 0.916265291 | 0.92202571 | 0.948648835 | 1 | 0.2106406 |
| 0.2 | 0.907244414 | 0.886289777 | 0.947192507 | 0.333333333 | 0.37455575 |
| 0.3 | 0.908955667 | 0.868659061 | 0.948285803 | 0.333333333 | 0.531855163 |
| 0.4 | 0.913161973 | 0.880561588 | 0.947893903 | 0.333333333 | 0.663554838 |
| 0.5 | 0.919091309 | 0.876642957 | 0.94764939 | 0.333333333 | 0.786467741 |
| **Average** | 0.912943731 | 0.886835819 | 0.947934088 | 0.466666667 |  |
| **Deviation** | 0.004410441 | 0.018505575 | 0.000503424 |  |  |
| Iteration-4 | | | | | |
| 0.1 | 0.912455595 | 0.92202571 | 0.948648835 | 1 | 0.210850629 |
| 0.2 | 0.908670729 | 0.886289777 | 0.947192507 | 0.333333333 | 0.375004399 |
| 0.3 | 0.909902014 | 0.868659061 | 0.948285803 | 0.333333333 | 0.513529106 |
| 0.4 | 0.913161973 | 0.880561588 | 0.947893903 | 0.333333333 | 0.665503748 |
| 0.5 | 0.920423919 | 0.876642957 | 0.94764939 | 0.333333333 | 0.781899148 |
| **Average** | 0.912922846 | 0.886835819 | 0.947934088 | 0.466666667 |  |
| **Deviation** | 0.004092616 | 0.018505575 | 0.000503424 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 0.914378672 | 0.92202571 | 0.948648835 | 1 | 0.210092985 |
| 0.2 | 0.909625367 | 0.886289777 | 0.947192507 | 0.333333333 | 0.378391257 |
| 0.3 | 0.910225515 | 0.868659061 | 0.948285803 | 0.333333333 | 0.515189547 |
| 0.4 | 0.911021378 | 0.880561588 | 0.947893903 | 0.333333333 | 0.665858928 |
| 0.5 | 0.919286139 | 0.876642957 | 0.94764939 | 0.333333333 | 0.787113957 |
| **Average** | 0.912907414 | 0.886835819 | 0.947934088 | 0.466666667 |  |
| **Deviation** | 0.003588155 | 0.018505575 | 0.000503424 |  |  |

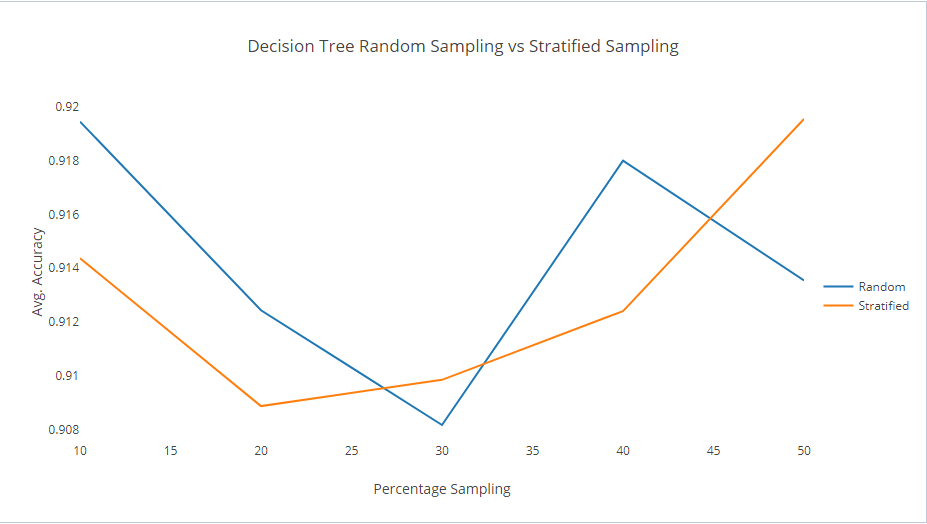
Tabel Rata – Rata *Random Sampling* West Nile

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Recommender | Time |
| 0.1 | 0.919450377 | 0.898132628 | 0.949533554 | 0.733333333 | 0.216728515 |
| 0.2 | 0.912432989 | 0.869692443 | 0.944414338 | 0.333333333 | 0.353831695 |
| 0.3 | 0.908160922 | 0.88664199 | 0.948082341 | 0.466666667 | 0.484915651 |
| 0.4 | 0.91799692 | 0.881252887 | 0.948550005 | 0.333333333 | 0.602905367 |
| 0.5 | 0.913538003 | 0.870927865 | 0.947116812 | 0.333333333 | 0.759189602 |
| **Average** | 0.914315842 | 0.881329563 | 0.94753941 | 0.44 | 0.483514166 |
| **Deviation** | 0.004048382 | 0.01052795 | 0.0017457 | 0.381333333 | 0.536871296 |

Tabel Rata – Rata *Stratified Sampling* West Nile

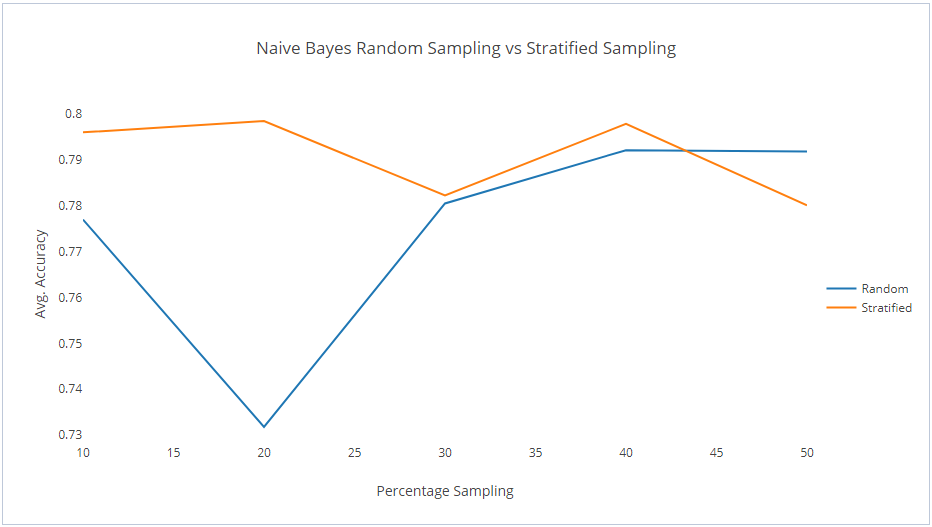
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Decision Tree | Naive Bayes | Logistic Regression | Recommender | Time |
| 0.1 | 0.914367717 | 0.92202571 | 0.948648835 | 1 | 0.209521397 |
| 0.2 | 0.908864365 | 0.886289777 | 0.947192507 | 0.333333333 | 0.374319843 |
| 0.3 | 0.909841743 | 0.868659061 | 0.948285803 | 0.333333333 | 0.517602207 |
| 0.4 | 0.912400633 | 0.880561588 | 0.947893903 | 0.333333333 | 0.667065001 |
| 0.5 | 0.919549904 | 0.876642957 | 0.94764939 | 0.333333333 | 0.789146438 |
| **Average** | 0.913004872 | 0.886835819 | 0.947934088 | 0.466666667 | 0.511530977 |
| **Deviation** | 0.003800239 | 0.018505575 | 0.000503424 | 0.36 | 0.571932893 |

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Decision Tree.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada Gambar 5.17



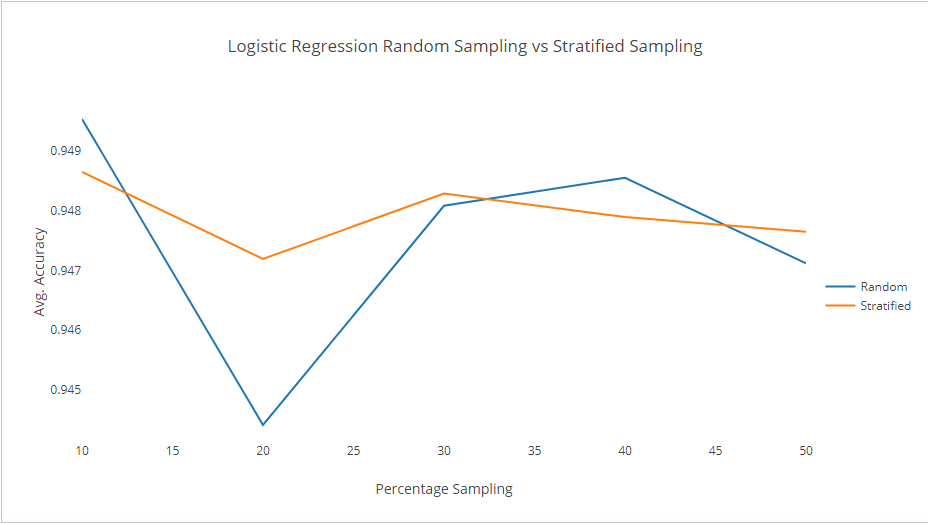
Grafik Perbandingan *Accuracy* *Decision Tree*

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *West Nile.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada Gambar 5.18



Grafik Perbandingan *Accuracy* *Naïve Bayes*

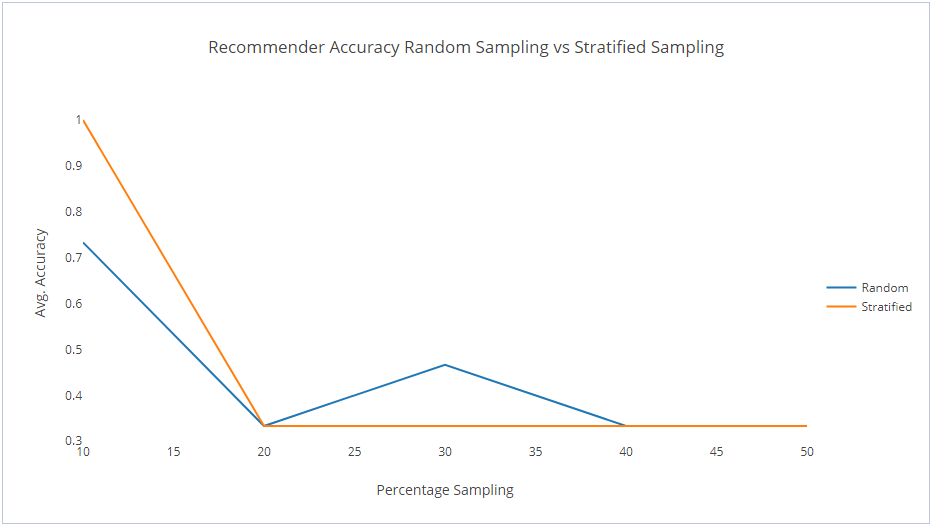
Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Logistic Regression.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada



Grafik Perbandingan *Accuracy* *Logistic Regression*

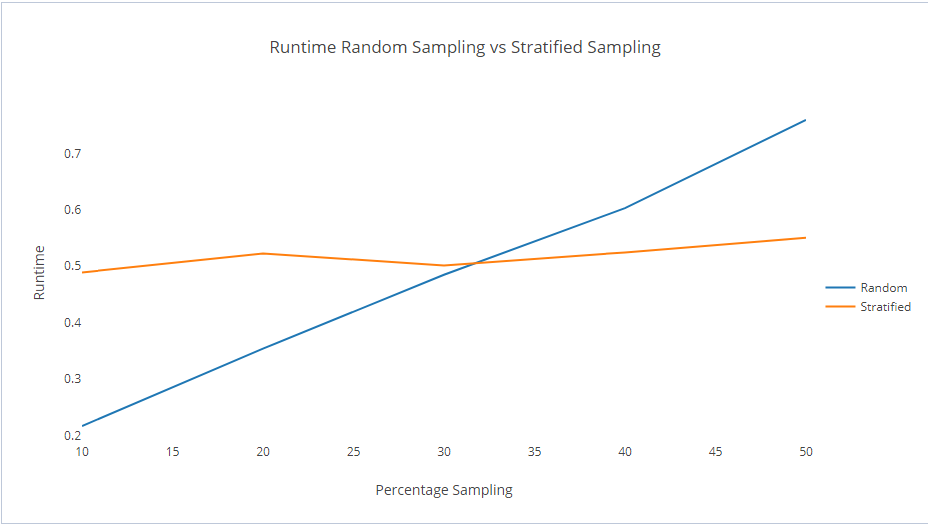
Berdasarkan Gambar 5.17, Gambar 5.18, Gambar 5.19 terlihat bahwa sistem rekomendasi yang menggunakan metode *Stratified Sampling* memiliki tingkat akurasi prediksi yang lebih konsisten. Pada percobaan *Human Resource* sistem rekomendasi juga memiliki tingkat *accuracy* prediksi yang lebih stabil dengan menggunakan *Stratified Sampling*.

Berdasarkan Tabel 5.7 dan Tabel 5.8 hasil *accuracy* sistem rekomendasi pada ukuran data sample dari 10 % sampai 50%, dapat dibentuk grafik yang membandingkan *accuracy* sistem rekomendasi pada *Random Sampling*  dan *Stratified Sampling.*



Grafik Perbandingan *Accuracy* Sistem Rekomendasi

Selain itu waktu yang diperlukan sistem rekomendasi untuk menjalankan proses rekomendasi pada data *human resource* dapat dibentuk dengan Gambar 5.21



*Runtime* Sistem Rekomendasi

Berdasarkab Gambar 5.20, sistem rekomendasi menggunakan *Stratified Sampling* lebih konsisten dibandingkan dengan menggunakan metode *Random Sampling.* Namun berdasarkan Gambar 5.21 kecepatan waktu proses sistem rekomendasi dengan menggunakan *Random Sampling*  memiliki waktu eksekusi yang lebih cepat dibandingkan dengan menggunakan *Stratified Sampling.*

### *House Price*

Type: *Regression*

Sampling: *Random Sampling & Stratified*

Data Size: 1460 Rows

All Data Result (menjalankan model dengan seluruh data):

1. *Lasso Regression*
2. *Regression Tree*
3. *SVR (Support Vector Regressor)*

Tabel *Iterasi Random Sampling House Price*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg. Tree | SVR | Lasso Regression | Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 78135.349 | 87082.94394 | 56124.62834 | 1 | 1.008792636 |
| 0.2 | 74503.122 | 73350.6601 | 57292.05318 | 0.333333333 | 4.894376569 |
| 0.3 | 81866.332 | 88050.02091 | 101167.8364 | 0 | 5.411463709 |
| 0.4 | 59431.940 | 79175.23141 | 52796.12506 | 1 | 8.176976914 |
| 0.5 | 61929.961 | 87171.6718 | 87311.92601 | 0 | 10.54848376 |
| **Average** | 71173.341 | 82966.10563 | 70938.5138 | 0.466666667 |  |
| **Deviation** | 8912.8854 | 5784.615097 | 19579.19268 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 66368.726 | 56055.23506 | 84745.61293 | 0.333333333 | 1.027567708 |
| 0.2 | 69306.189 | 73609.62696 | 79262.83984 | 0 | 4.587572282 |
| 0.3 | 68976.375 | 78467.80828 | 89642.78216 | 0 | 5.652771506 |
| 0.4 | 64985.061 | 85315.1085 | 93296.40061 | 0 | 8.342002689 |
| 0.5 | 57012.054 | 78258.78199 | 40619.62707 | 1 | 11.02239175 |
| **Average** | 65329.681 | 74341.31216 | 77513.45252 | 0.266666667 |  |
| **Deviation** | 4460.7686 | 9876.383196 | 19040.69822 |  |  |
| Iteration-3 | | | | | |
| 0.1 | 88348.250 | 85028.51208 | 52881.80992 | 0.333333333 | 1.001001396 |
| 0.2 | 65155.473 | 75326.53115 | 89336.39398 | 0 | 4.882220533 |
| 0.3 | 70963.163 | 85904.63201 | 90882.3081 | 0 | 5.37597024 |
| 0.4 | 68801.014 | 85777.76709 | 45684.17219 | 1 | 8.518256046 |
| 0.5 | 58213.899 | 84025.16731 | 47832.83295 | 1 | 10.745438 |
| **Average** | 70296.360 | 83212.52193 | 65323.50343 | 0.466666667 |  |
| **Deviation** | 10009.655 | 3999.37502 | 20377.89995 |  |  |

Tabel *Iterasi Random Sampling House Price* (Lanjutan)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Iteration-4 | | | | | |
| 0.1 | 63882.477 | 86624.64638 | 76540.39968 | 0.333333333 | 0.974041416 |
| 0.2 | 66015.980 | 69855.85767 | 57567.04652 | 1 | 4.596278414 |
| 0.3 | 63600.661 | 75159.08813 | 58076.70074 | 1 | 5.305688529 |
| 0.4 | 62158.699 | 85394.17429 | 73018.74211 | 0.333333333 | 8.606855749 |
| 0.5 | 61633.671 | 86608.7009 | 58681.78572 | 1 | 10.54293466 |
| **Average** | 63458.298 | 80728.49347 | 64776.93495 | 0.733333333 |  |
| **Deviation** | 1533.6331 | 6933.145335 | 8250.248016 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 67690.168 | 87654.72918 | 72015.45207 | 0.333333333 | 1.131018271 |
| 0.2 | 69529.308 | 81348.23924 | 73277.76956 | 0.333333333 | 4.644047413 |
| 0.3 | 67141.530 | 77048.967 | 81229.19553 | 0 | 5.356164449 |
| 0.4 | 54673.562 | 79955.07588 | 58474.44796 | 0.333333333 | 8.552428495 |
| 0.5 | 69856.548 | 84498.79326 | 52235.53196 | 1 | 10.94946213 |
| **Average** | 65778.224 | 82101.16091 | 67446.47942 | 0.4 |  |
| **Deviation** | 5648.55 | 3668.460078 | 10551.90021 |  |  |

Tabel *Iterasi Stratified Sampling House Price*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg.Tree | SVR | Lasso Regression | Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 62882.91972 | 79033.16995 | 67695.16962 | 0.333333333 | 5.047679562 |
| 0.2 | 59828.50243 | 74341.17566 | 81008.96363 | 0 | 5.173338607 |
| 0.3 | 59510.45901 | 78687.20259 | 76652.85427 | 0.333333333 | 5.130644764 |
| 0.4 | 59855.71634 | 74101.95722 | 88510.41106 | 0 | 5.373038628 |
| 0.5 | 61024.68442 | 76040.70798 | 69169.96328 | 0.333333333 | 5.222469313 |
| **Average** | 60620.45638 | 76440.84268 | 76607.47237 | 0.2 |  |
| **Deviation** | 1243.081107 | 2088.319332 | 7691.564419 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 54204.06105 | 83648.20268 | 91228.58978 | 0 | 5.330628489 |
| 0.2 | 71459.08296 | 83223.45138 | 86862.98005 | 0 | 5.214137788 |
| 0.3 | 71100.99863 | 83086.21452 | 69324.26872 | 1 | 5.199185028 |
| 0.4 | 70204.49814 | 84646.63772 | 101453.7105 | 0 | 5.353344999 |
| 0.5 | 72534.46068 | 79156.70439 | 66745.17959 | 1 | 5.692859227 |
| **Average** | 67900.62029 | 82752.24214 | 83122.94572 | 0.4 |  |
| **Deviation** | 6888.848132 | 1878.986286 | 13223.77381 |  |  |
| Iteration-3 | | | | | |
| 0.1 | 82858.54969 | 89782.85789 | 84468.3237 | 0.333333333 | 5.291162059 |
| 0.2 | 64697.77314 | 75307.10877 | 92102.74052 | 0 | 5.080075742 |
| 0.3 | 80593.76205 | 85251.08651 | 100155.2174 | 0 | 5.114358887 |
| 0.4 | 68370.71795 | 81930.15659 | 103695.3301 | 0 | 5.285394503 |
| 0.5 | 64949.2075 | 76332.42539 | 61580.45682 | 1 | 5.163137345 |
| **Average** | 72294.00207 | 81720.72703 | 88400.41371 | 0.266666667 |  |
| **Deviation** | 7842.659519 | 5434.584558 | 14970.56124 |  |  |
| Iteration-4 | | | | | |
| 0.1 | 67722.9902 | 90563.95423 | 85688.90869 | 0.333333333 | 5.241400898 |
| 0.2 | 65627.32447 | 81002.01306 | 79305.37028 | 0.333333333 | 5.297909018 |
| 0.3 | 65360.77384 | 86768.9219 | 95329.92656 | 0 | 5.251621953 |
| 0.4 | 58062.2681 | 73975.61344 | 76795.21966 | 0 | 5.257370082 |
| 0.5 | 63224.57441 | 87558.10197 | 90143.12752 | 0 | 5.02048133 |
| **Average** | 63999.58621 | 83973.72092 | 85452.51054 | 0.133333333 |  |
| **Deviation** | 3292.977089 | 5880.226388 | 6816.995494 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 71112.70579 | 86126.91795 | 107497.3249 | 0 | 5.021712182 |
| 0.2 | 62780.09428 | 87074.1264 | 107680.568 | 0 | 5.110568831 |
| 0.3 | 73908.55515 | 91881.4572 | 112290.3502 | 0 | 5.181834343 |
| 0.4 | 62965.39919 | 83055.05529 | 69163.60139 | 0.333333333 | 5.078266119 |
| 0.5 | 71001.40479 | 85279.62641 | 90447.59585 | 0 | 5.064206625 |
| **Average** | 68353.63184 | 86683.43665 | 97415.88806 | 0.066666667 |  |
| **Deviation** | 4595.165625 | 2919.506936 | 15968.70342 |  |  |

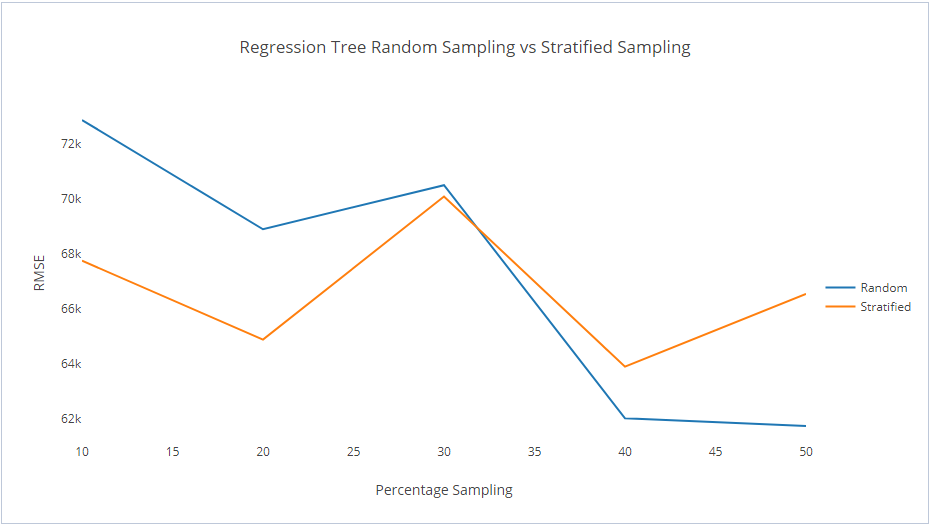
Tabel Rata – Rata *Random Sampling House Price*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg.Tree | SVR | Lasso Regression | Recommender | Time |
| 0.1 | 72884.99484 | 80489.21333 | 68461.58059 | 0.466666667 | 1.028484286 |
| 0.2 | 68902.01475 | 74698.18302 | 71347.22062 | 0.333333333 | 4.720899042 |
| 0.3 | 70509.6128 | 80926.10327 | 84199.76459 | 0.2 | 5.420411687 |
| 0.4 | 62010.05568 | 83123.47143 | 64653.97759 | 0.533333333 | 8.439303979 |
| 0.5 | 61729.22719 | 84112.62305 | 57336.34074 | 0.8 | 10.76174206 |
| **Average** | 67207.18105 | 80669.91882 | 69199.77682 | 0.466666667 | 6.074168211 |
| **Deviation** | 6113.100049 | 6052.395745 | 15559.98782 | 0.466666667 | 7.083304996 |

Tabel Rata – Rata *Stratified Sampling House Price*

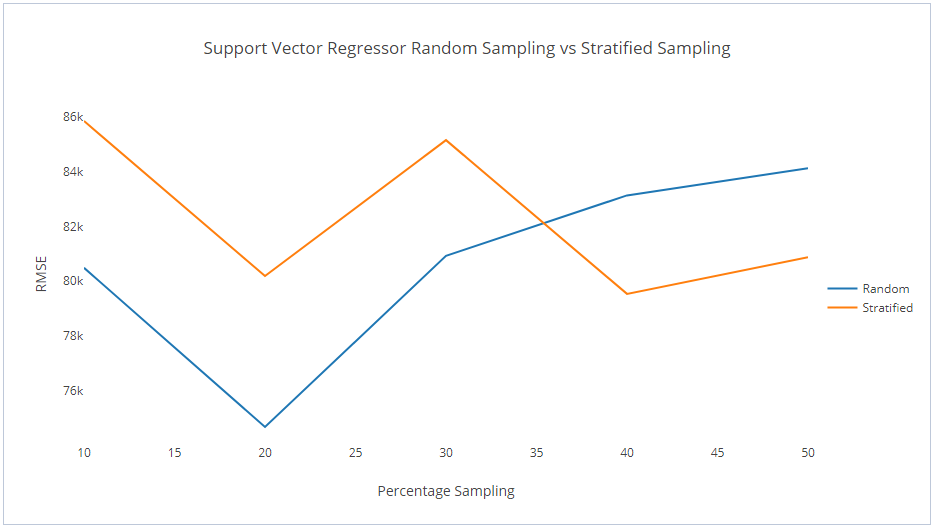
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg.Tree | SVR | Lasso Regression | Recommender | Time |
| 0.1 | 67756.24529 | 85831.02054 | 87315.66335 | 0.2 | 5.186516638 |
| 0.2 | 64878.55546 | 80189.57506 | 89392.12449 | 0.066666667 | 5.175205997 |
| 0.3 | 70094.90974 | 85134.97655 | 90750.52342 | 0.266666667 | 5.175528995 |
| 0.4 | 63891.71995 | 79541.88405 | 87923.65453 | 0.066666667 | 5.269482866 |
| 0.5 | 66546.86636 | 80873.51323 | 75617.26461 | 0.466666667 | 5.232630768 |
| **Average** | 66633.65936 | 82314.19388 | 86199.84608 | 0.213333333 | 5.207873053 |
| **Deviation** | 4772.546294 | 3640.3247 | 11734.31968 | 0.216 | 5.212144336 |

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Regression Tree.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada Gambar 5.22



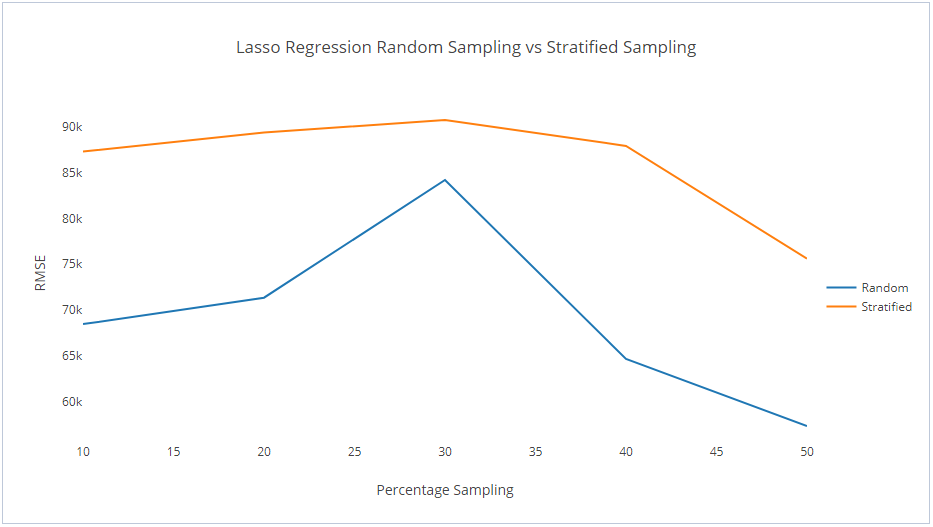
Perbandingan *Accuracy Regression Tree*

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Support Vector Regressor.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada Gambar 5.23



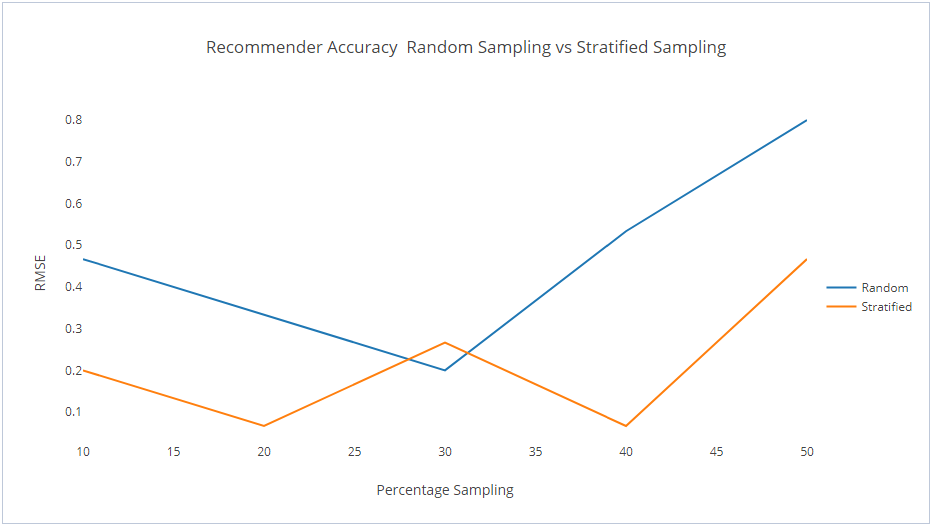
Perbandingan *Accuracy Support Vector Regressor*

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Lasso Regression.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada



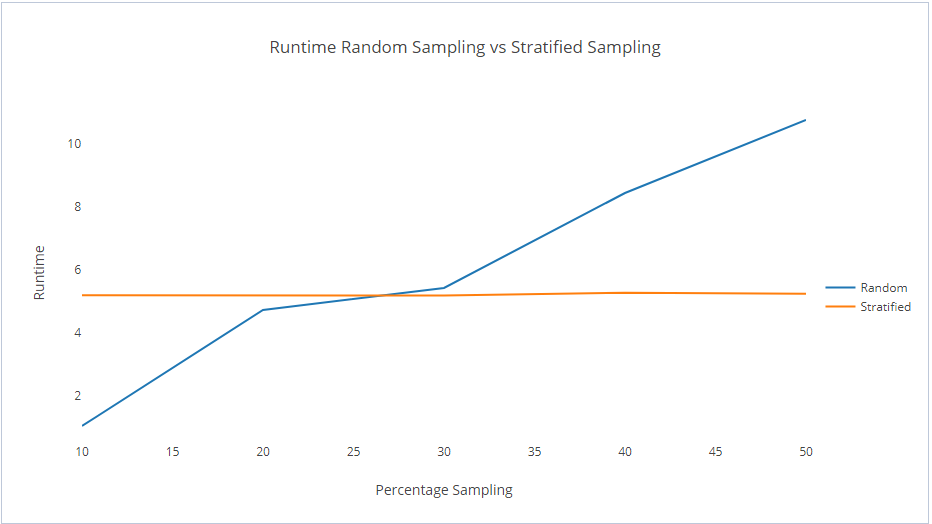
Perbandingan *Accuracy Lasso Regression*

Berdasarkan Tabel 5.12 dan Tabel 5.13 hasil *accuracy* sistem rekomendasi pada ukuran data sample dari 10 % sampai 50%, dapat dibentuk grafik yang membandingkan *accuracy* sistem rekomendasi pada *Random Sampling*  dan *Stratified Sampling.*



Grafik Perbandingan *Accuracy* Sistem Rekomendasi

Selain itu waktu yang diperlukan sistem rekomendasi untuk menjalankan proses rekomendasi pada data *human resource* dapat dibentuk dengan Gambar 5.26



*Runtime* Sistem Rekomendasi

### *Bakery Demand*

Type: *Regression*

Sampling: *Random Sampling & Stratified*

Data Size: 25127 Rows

All Data Result (menjalankan model dengan seluruh data):

1. *Lasso Regression*
2. *SVR (Support Vector Regressor)*
3. *Regression Tree*

Tabel *Iterasi Random Sampling* *Bakery Demand*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg.Tree | SVR | Lasso Regression | Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 14.94511545 | 12.863497 | 10.75600347 | 1 | 4.609953038 |
| 0.2 | 14.81509747 | 13.16081326 | 10.75989959 | 1 | 17.2658101 |
| 0.3 | 15.08074219 | 13.25574818 | 10.93781173 | 1 | 37.07867372 |
| 0.4 | 14.92979123 | 13.11738809 | 10.82049279 | 1 | 63.89736976 |
| 0.5 | 14.85611178 | 13.17482339 | 10.90565383 | 1 | 98.59590051 |
| **Average** | 14.92537162 | 13.11445398 | 10.83597228 | 1 |  |
| **Deviation** | 0.091111669 | 0.13322591 | 0.074361912 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 14.55506727 | 13.10733481 | 10.70480676 | 1 | 4.962709736 |
| 0.2 | 15.0050859 | 13.21334941 | 10.89623602 | 1 | 17.15379746 |
| 0.3 | 14.97777494 | 13.16767027 | 10.87013461 | 1 | 37.57189853 |
| 0.4 | 14.81087527 | 13.05215865 | 10.83481107 | 1 | 63.72743707 |
| 0.5 | 14.77615519 | 13.17098644 | 10.8590717 | 1 | 98.80345775 |
| **Average** | 14.82499171 | 13.14229991 | 10.83301203 | 1 |  |
| **Deviation** | 0.162000071 | 0.056325143 | 0.067073531 |  |  |
| Iteration-3 | | | | | |
| 0.1 | 14.78613594 | 13.0217625 | 10.96140224 | 1 | 4.608635683 |
| 0.2 | 14.61859236 | 12.94547452 | 10.60195086 | 1 | 16.82411062 |
| 0.3 | 14.70640754 | 13.09087591 | 10.83035268 | 1 | 36.77519141 |
| 0.4 | 14.97223041 | 13.10763913 | 10.88519196 | 1 | 63.57843363 |
| 0.5 | 14.87979099 | 13.14491814 | 10.85565613 | 1 | 99.26514043 |
| **Average** | 14.79263145 | 13.06213404 | 10.82691078 | 1 |  |
| **Deviation** | 0.124594488 | 0.070695573 | 0.12077231 |  |  |

Tabel *Iterasi Random Sampling Bakery Demand* (Lanjutan)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Iteration-4 | | | | | |
| 0.1 | 14.73276496 | 13.10109008 | 10.76464527 | 1 | 4.651951915 |
| 0.2 | 14.89080217 | 13.13295304 | 10.75959044 | 1 | 16.34429028 |
| 0.3 | 14.98662258 | 13.22547943 | 10.93173606 | 1 | 35.55187392 |
| 0.4 | 15.14965806 | 13.1772501 | 10.91661756 | 1 | 61.03398844 |
| 0.5 | 14.69273961 | 13.19181078 | 10.87422721 | 1 | 94.3098312 |
| **Average** | 14.89051747 | 13.16571669 | 10.84936331 | 1 |  |
| **Deviation** | 0.167566141 | 0.043898887 | 0.073706082 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 15.22663354 | 13.35570404 | 10.8020697 | 1 | 4.492430885 |
| 0.2 | 15.14277717 | 13.36710963 | 11.14338873 | 1 | 16.28968489 |
| 0.3 | 14.68228719 | 13.14922231 | 10.80420372 | 1 | 35.5545889 |
| 0.4 | 15.06165925 | 13.28587805 | 10.95201186 | 1 | 60.74304411 |
| 0.5 | 14.56468075 | 13.13134895 | 10.81787796 | 1 | 94.92105101 |
| **Average** | 14.93560758 | 13.2578526 | 10.90391039 | 1 |  |
| **Deviation** | 0.262778349 | 0.100101093 | 0.132195426 |  |  |

Tabel *Iterasi Stratified Sampling Bakery Demand*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg.Tree | SVR | Lasso Regression | Recommender | Time |
| Iteration-1 | | | | | |
| 0.1 | 14.68643406 | 13.02145478 | 10.80175117 | 1 | 35.39479736 |
| 0.2 | 14.6730025 | 13.16028575 | 10.74814123 | 1 | 35.28308345 |
| 0.3 | 14.79227287 | 13.19887906 | 10.80398659 | 1 | 35.43928946 |
| 0.4 | 14.80641893 | 13.08650159 | 10.76140599 | 1 | 35.63008799 |
| 0.5 | 14.57950879 | 13.02557282 | 10.79846381 | 1 | 35.36424233 |
| **Average** | 14.70752743 | 13.0985388 | 10.78274976 | 1 |  |
| **Deviation** | 0.083650453 | 0.071122046 | 0.02329075 |  |  |
| Iteration-2 | | | | | |
| 0.1 | 15.15649831 | 13.13470904 | 10.85842412 | 1 | 35.40547146 |
| 0.2 | 14.51546313 | 13.07523566 | 10.73475922 | 1 | 35.31398523 |
| 0.3 | 14.91811662 | 13.13614297 | 10.77359406 | 1 | 35.39028082 |
| 0.4 | 15.09823731 | 13.19054937 | 10.93578485 | 1 | 35.4644959 |
| 0.5 | 14.64188575 | 13.13755425 | 10.77493709 | 1 | 35.39141123 |
| **Average** | 14.86604022 | 13.13483826 | 10.81549987 | 1 |  |
| **Deviation** | 0.250656283 | 0.036511077 | 0.072431284 |  |  |
| Iteration-3 | | | | | |
| 0.1 | 15.10888507 | 13.23893849 | 11.03482605 | 1 | 35.80830968 |
| 0.2 | 14.85747629 | 13.02716448 | 10.75257726 | 1 | 36.24404441 |
| 0.3 | 14.87039621 | 13.01307172 | 10.79573365 | 1 | 36.57878991 |
| 0.4 | 14.69587223 | 13.06544611 | 10.69132245 | 1 | 35.81897205 |
| 0.5 | 14.92160493 | 13.06120823 | 10.74954636 | 1 | 35.81645537 |
| **Average** | 14.89084695 | 13.0811658 | 10.80480116 | 1 |  |
| **Deviation** | 0.132690353 | 0.081351808 | 0.119709245 |  |  |
| Iteration-4 | | | | | |
| 0.1 | 15.05666456 | 13.2248274 | 10.94358996 | 1 | 35.9157343 |
| 0.2 | 14.94011842 | 13.11038983 | 10.78725954 | 1 | 35.609272 |
| 0.3 | 14.77654161 | 13.09303875 | 10.76913841 | 1 | 35.17394963 |
| 0.4 | 14.80631149 | 13.17897801 | 10.87408057 | 1 | 35.73008168 |
| 0.5 | 14.89503505 | 13.28444611 | 10.91172962 | 1 | 36.09137004 |
| **Average** | 14.89493423 | 13.17833602 | 10.85715962 | 1 |  |
| **Deviation** | 0.100069102 | 0.071152682 | 0.06836397 |  |  |
| Iteration-5 | | | | | |
| 0.1 | 15.18981035 | 13.36959879 | 11.11449147 | 1 | 35.16623097 |
| 0.2 | 14.95139126 | 13.08478519 | 10.78545031 | 1 | 36.13648856 |
| 0.3 | 14.78588688 | 13.23755839 | 10.88286599 | 1 | 35.94717197 |
| 0.4 | 14.97805922 | 13.20760105 | 11.02452511 | 1 | 36.54516361 |
| 0.5 | 14.67035445 | 13.05111675 | 10.74140165 | 1 | 35.81329283 |
| **Average** | 14.91510043 | 13.19013203 | 10.9097469 | 1 |  |
| **Deviation** | 0.177393975 | 0.114180514 | 0.141142996 |  |  |

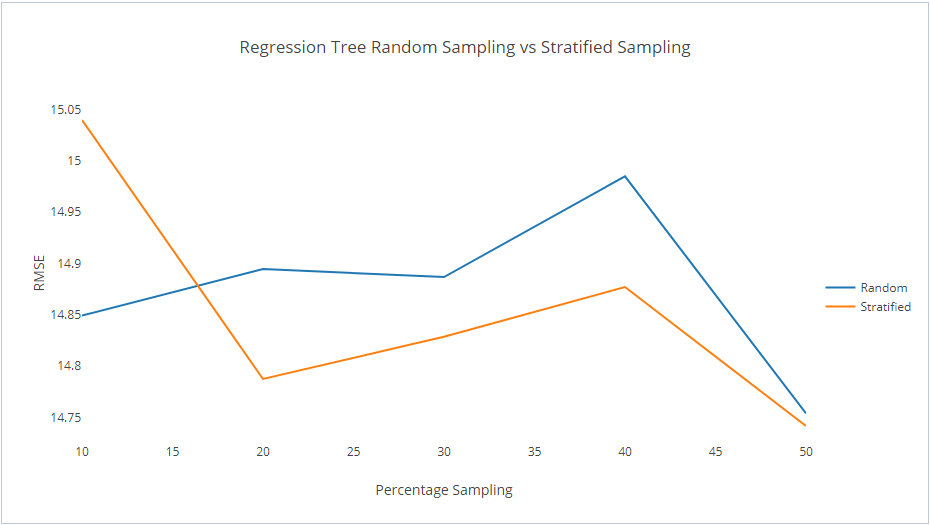
Tabel Rata – Rata *Random Sampling Bakery Demand*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg.Tree | SVR | Lasso Regression | Recommender | Time |
| 0.1 | 14.84914343 | 13.08987768 | 10.79778549 | 1 | 4.665136251 |
| 0.2 | 14.89447101 | 13.16393997 | 10.83221313 | 1 | 16.77553867 |
| 0.3 | 14.88676689 | 13.17779922 | 10.87484776 | 1 | 36.50644529 |
| 0.4 | 14.98484284 | 13.14806281 | 10.88182505 | 1 | 62.5960546 |
| 0.5 | 14.75389567 | 13.16277754 | 10.86249737 | 1 | 97.17907618 |
| **Average** | 14.87382397 | 13.14849144 | 10.84983376 | 1 | 43.5444502 |
| **Deviation** | 0.161610144 | 0.080849321 | 0.093621852 | 0 | 33.21850666 |

Tabel Rata – Rata *Stratified Sampling Bakery Demand*

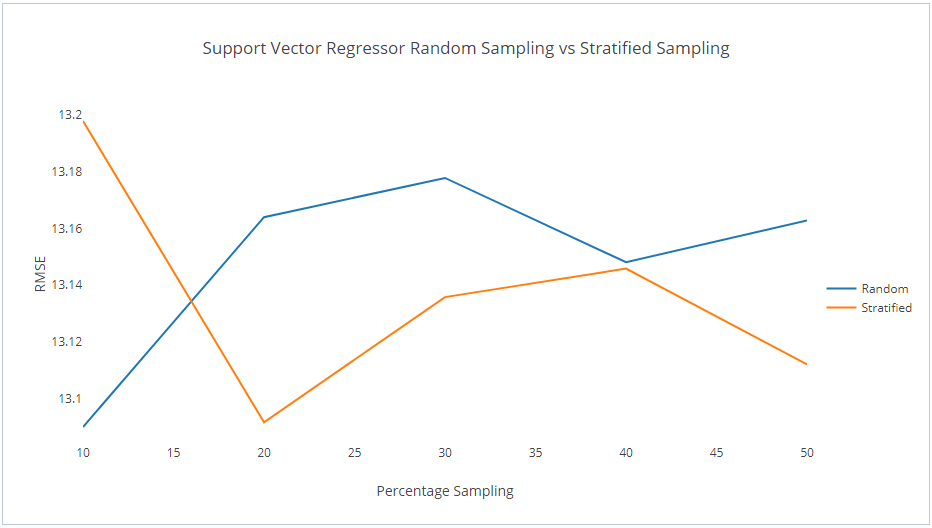
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sampling | Reg.Tree | SVR | Lasso Regression | Recommender | Time |
| 0.1 | 15.03965847 | 13.1979057 | 10.95061656 | 1 | 35.53810875 |
| 0.2 | 14.78749032 | 13.09157218 | 10.76163751 | 1 | 35.71737473 |
| 0.3 | 14.82864284 | 13.13573818 | 10.80506374 | 1 | 35.70589636 |
| 0.4 | 14.87697984 | 13.14581523 | 10.85742379 | 1 | 35.83776025 |
| 0.5 | 14.7416778 | 13.11197963 | 10.79521571 | 1 | 35.69535436 |
| **Average** | 14.85488985 | 13.13660218 | 10.83399146 | 1 | 35.69889889 |
| **Deviation** | 0.148892033 | 0.074863625 | 0.084987649 | 0 | 0.095434615 |

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Regression Tree.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada Gambar 5.27



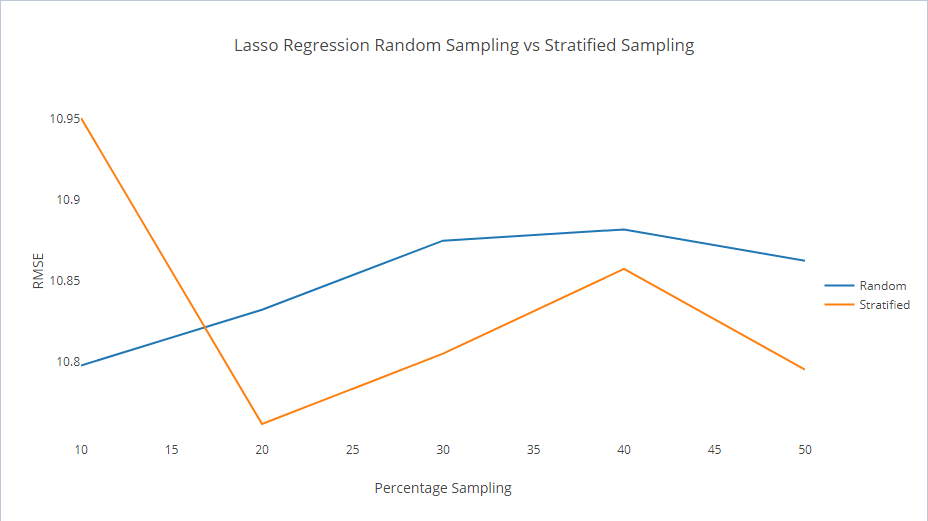
Perbandingan *Accuracy Regression Tree*

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Support Vector Regressor.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada Gambar 5.28



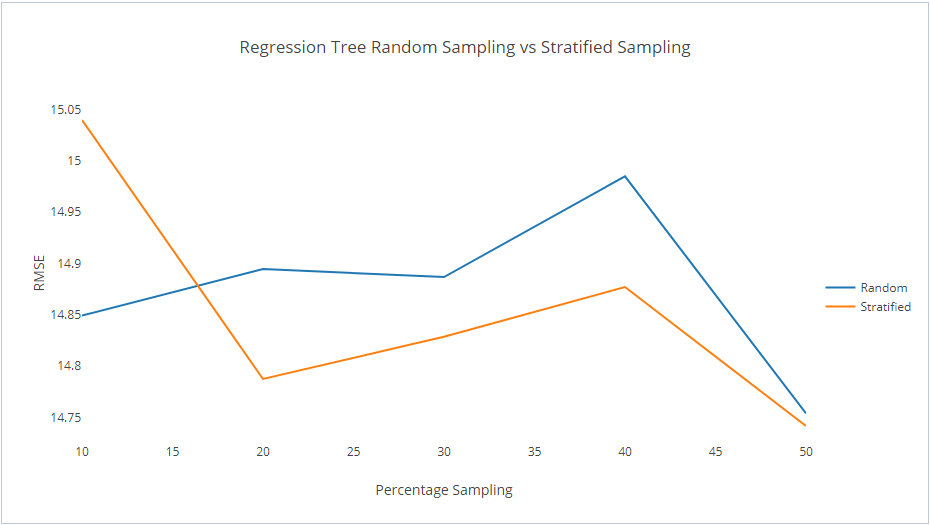
Perbandingan *Accuracy Support Vector Regressor*

Berdasarkan rata – rata dari setiap ukuran sampling (10 % - 50%), terlihat perbandingan *accuracy* antara *Random Sampling* dan *Stratified Sampling* pada metode *Lasso Regression.* Perbandingan tersebut dapat dibentuk dalam bentuk grafikpada Gambar 5.29



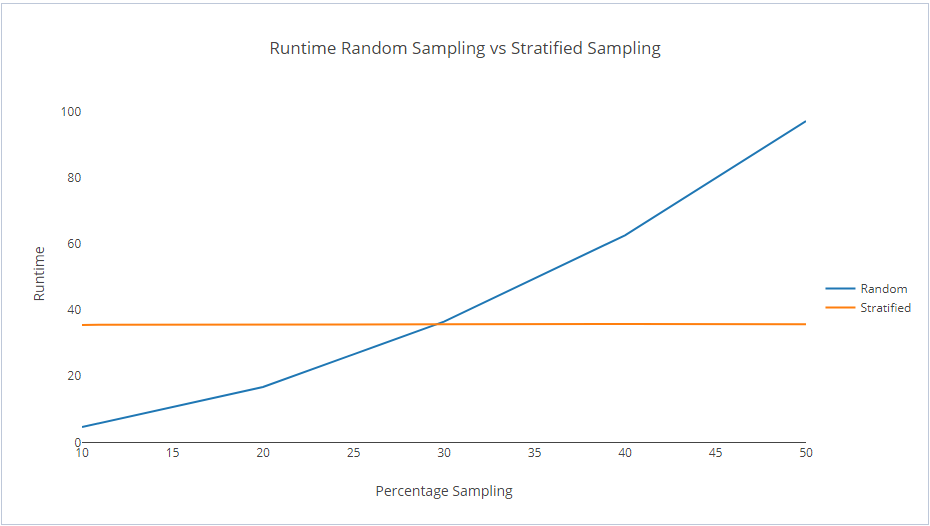
Perbandingan *Accuracy Lasso Regression*

Berdasarkan Tabel 5.17 dan Tabel 5.18 hasil *accuracy* sistem rekomendasi pada ukuran data sample dari 10 % sampai 50%, dapat dibentuk grafik yang membandingkan *accuracy* sistem rekomendasi pada *Random Sampling*  dan *Stratified Sampling.*



Grafik Perbandingan *Accuracy* Sistem Rekomendasi

Selain itu waktu yang diperlukan sistem rekomendasi untuk menjalankan proses rekomendasi pada data *human resource* dapat dibentuk dengan



*Runtime* Sistem Rekomendasi

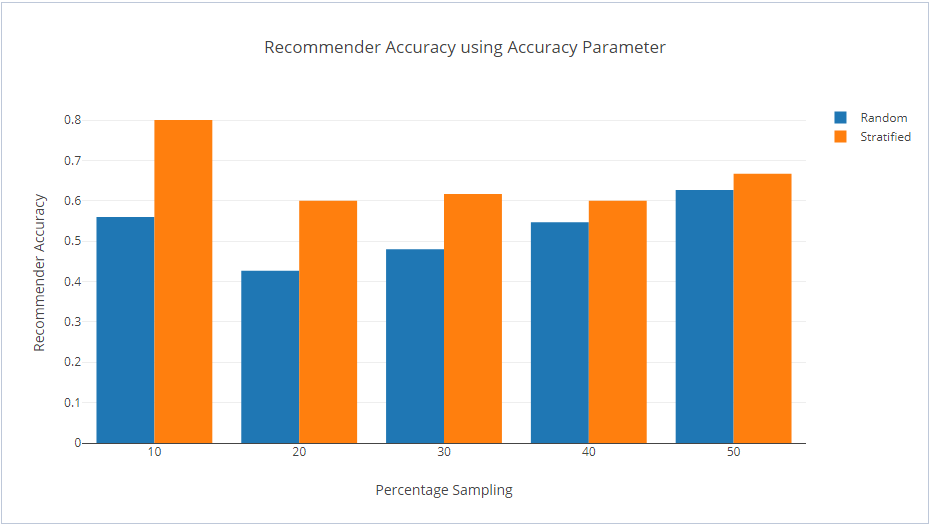
### Overall

Secara keseluruhan dari pengujian yang telah dilakukan, dapat dibentuk sebuah tabel yang memberikan hasil rata – rata dari semua pengujian yang pernah dilakukan.

Rata – Rata *Accuracy* Recommender (Parameter *Accuracy*)

|  |  |  |
| --- | --- | --- |
|  | Random Sampling | Stratified Sampling |
| 10 | 0.56 | 0.8 |
| 20 | 0.426666667 | 0.6 |
| 30 | 0.48 | 0.616666667 |
| 40 | 0.546666667 | 0.6 |
| 50 | 0.626666667 | 0.666666667 |
| **AVG** | 0.528 | 0.656666667 |

Pada Tabel 5.19 dapat terlihat bahwa rata – rata *accuracy recommender* lebih baik ketika *sampling* yang digunakan adalah *Stratified Sampling*.



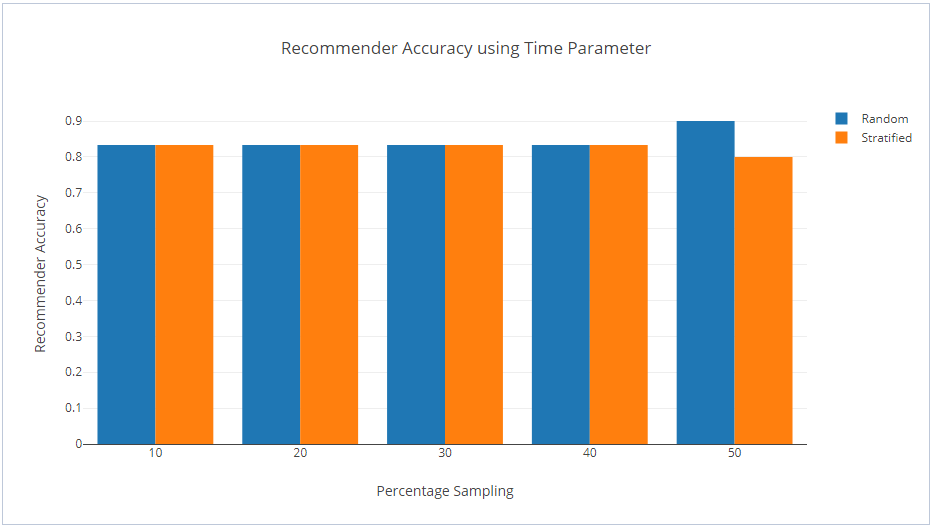
Grafik Keseluruhan *accuracy* Recommender (*Accuracy*)

Selanjutnya terdapat tabel *accuracy recommender* ketika parameter yang digunakan adalah *time.*

Rata – Rata *Accuracy* Recommender (Parameter *Time*)

|  |  |  |
| --- | --- | --- |
|  | Random Sampling | Stratified Sampling |
| 10 | 0.833333333 | 0.833333333 |
| 20 | 0.833333333 | 0.833333333 |
| 30 | 0.833333333 | 0.833333333 |
| 40 | 0.833333333 | 0.833333333 |
| 50 | 0.9 | 0.8 |
| **AVG** | 0.846666667 | 0.826666667 |

Pada Tabel 5.20 dapat terlihat bahwa rata – rata *accuracy recommender* lebih baik ketika *sampling* yang digunakan adalah *Random Sampling*.



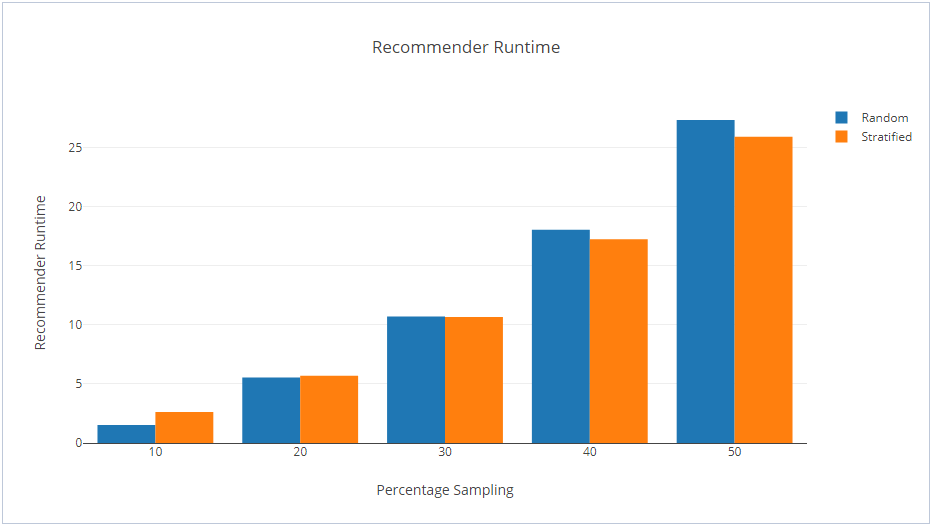
Grafik Keseluruhan *accuracy* Recommender (*Time*)

Selanjutnya, performa kerja *recommender*, terhadap waktu (lama recommender berjalan sampai memberikan rekomendasi) dapat dilihat pada Tabel 5.21

Lama *Running* *Recommender*

|  |  |  |
| --- | --- | --- |
|  | Random Sampling | Stratified Sampling |
| 10 | 1.521333689 | 2.61258929 |
| 20 | 5.530719319 | 5.678675401 |
| 30 | 10.69425187 | 10.64988582 |
| 40 | 18.02580762 | 17.23231939 |
| 50 | 27.31614276 | 25.89492924 |
| **AVG** | 12.61765105 | 12.41367983 |

Pada Tabel 5.21 terlihat bahwa rata – rata jalan *recommender* hampir sama, tetapi jika dilihat lebih lanjut, *Random Sampling*  memiliki *running* waktu lebih cepat pada 10% - 20%, dan lebih lambat setelahnya. *Stratified Sampling* memiliki *running* waktu lebih lama pada 10% - 20% tetapi lebih cepat pada setelahnya.



Grafik Lama *Running* *Recommender*