



Nama: Yap Shen Yuan No. telefon: 0177224115

Email: yuanyap_0208@hotmail.com & zatur@ums.edu.my

Fakulti Sains dan Sumber Alam, University Malaysia Sabah, 88400, Kota Kinabalu, Sabah.

ABSTRAK

Ramalan cuaca mempunyai banyak cabaran dalam sejarah. Long Short Term Memory (LSTM) adalah salah satu mesin ramalan cuaca. Untuk menguji hipotesis tentang hubungan antara purata populasi dengan purata ramalan populasi dalam ramalan faktor cuaca. Dengan ujian hipotesis dapat membuktikan bahawa teknik LSTM adalah satu mesin ramalan cuaca yang dapat mengukur ketidakpastian ini.

PERNYATAAN MASALAH

Aktiviti-aktiviti manusia sangat dipengaruhi oleh keadaan cuaca dan sebarang kaedah yang dapat memantau keadaan cuaca sangat membantu dalam mengawal aktiviti-aktiviti tersebut. Kini, manusia banyak bergantung dengan bantuan teknologi untuk menjalankan kebanyakan perkara dengan pantas. Oleh kerana itu, terdapat banyak bahan atau aplikasi yang dicipta oleh orang ramai. Untuk menyelesaikan masalah ini, ujian hipotesis telah dijalankan dalam kajian ini untuk mengumpulkan bukti dengan menjalankan sampel ramalan faktor cuaca.

METODOLOGI

- Langkah-langkah berikut digunakan untuk mencapai objektif:
 - Persediaan
 - Pengumpulan data
 - Pemprosesan data
 - Model latihan

2. Memahami proses LSTM

3. Ujian hipotesis dengan ujian t-test

Algoritma Ujian T-test

Langkah pertama : Menentukan hipotesis nol, H_0 dan hipotesis alternatif, H_1 .

Langkah kedua : Mencari nilai kritikal dengan menggunakan tahap kebebasan, df dan nilai alpha.

Langkah ketiga : Mencari nilai t-test dengan menggunakan formula:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (1)$$

di mana x_1 adalah nilai sebenar dan x_2 adalah nilai ramalan.

Langkah keempat : Melukis graph dan buat keputusan untuk menerima atau menolak hipotesis nol. Jika hitung nilai ujian berada dalam kawasan kritikal, hipotesis nol akan ditolak.

Langkah ketiga : Kesimpulan ujian hipotesis.

PENUTUP

Dalam kajian ini, model yang dicadangkan untuk sistem peramalan faktor cuaca dilaksanakan dengan menggunakan RNN dengan LSTM. Dalam model ini, data dilatih menggunakan algoritma LSTM. Dari graf keputusan diperhatikan bahawa LSTM memberikan hasil yang besar dengan perbezaan yang rendah di antara teknik ramalan faktor-faktor meteorologi cuaca yang lain. Untuk menentukan ketidakpastian data-data ramalan, ujian hipotesis telah dijalankan. Keempat-empat set data (suhu, kelembapan, kelajuan angin dan tekanan) terdapat tiga set data telah menyokong nilai perbezaan antara nilai purata populasi dan nilai purata ramalan populasi adalah kosong. Ini bermaksud teknik LSTM adalah teknik yang dapat menentukan ketidakpastian semasa ramalan set data-data faktor cuaca.

PENDAHULUAN

Nilai statistik adalah apa yang digunakan untuk menentukan kebenaran pernyataan asal. Proses ini mengandungi ketidakpastian, namun, kajian ini adalah untuk mengukur ketidakpastian ramalan faktor cuaca.

OBJEKTIF

Untuk mengaplikasikan model *Long Short Term Memory* mempunyai kepastian dengan menggunakan ujian hipotesis (ujian t-test) bagi kawasan Kota Kinabalu.

- Mengaplikasikan teknik LSTM sebagai model meramal faktor-faktor cuaca.
- Menjalankan ujian hipotesis bagi menentukan terdapat bukti yang mencukupi. (Tuntutan: H_0 : purata populasi – purata ramalan populasi = 0)

KEPUTUSAN & RINGKASAN

Sistem menggunakan 80% (292 data) sebelum data untuk latihan data yang seterusnya 20% data (73 data). Untuk memastikan ketepatan, ujian hipotesis telah dijalankan. Jadual keputusan ujian hipotesis yang ditunjukkan dengan nilai kritikal=0.05.

Faktor Cuaca	\bar{X}_1	\bar{X}_2	s_1	s_2	T-test	P-value
Suhu	28.0959	28.3788	0.6643	0.3319	6.1203	<0.00001
Kelembapan	81.8493	79.8930	13.1228	29.1458	3.5307	0.0007
Kelajuan angin	5.4452	5.9313	3.4831	3.8656	1.6494	0.1034
Tekanan	29.7664	29.7747	0.0013	0.0025	598.7660	<0.00001

Jika p-value lebih kecil dari nilai kritikal (0.05), tidak menerima dakwaan. Dari jadual atas, dapat satu keputusan, ketiga-tiga faktor cuaca (suhu, kelembapan dan tekanan) mempunyai bukti yang cukup untuk menyokong dakwaan bahawa tidak ada perbezaan yang signifikan antara nilai purata dan nilai purata ramalan.

RUJUKAN

- Olah, C. (2015). Understanding LSTM Networks [Blog]. *Web Page*, 1–13. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Crouse, C. F., & Lehmann, E. L. (1960). Testing Statistical Hypothesis. In *Biometrika* (Vol. 47, Issue 1/2). <https://doi.org/10.2307/2332982>
- Tomáš Mikolov, Martin Karafiát, Lukáš Burget, Jan Černocký, S. K. (2010). Recurrent Neural Network Based Language Model. *Journal of Quality Technology*, 25(4), 271–287. <https://doi.org/10.1080/00224065.1993.11979473>