

DAFTAR ISI

HALAMAN BEBAS PLAGIASI	iv
PRAKATA	v
DAFTAR ISI	vi
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiv
INTISARI	xv
ABSTRACT	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat	3
BAB II TINJAUAN PUSTAKA	4
BAB III LANDASAN TEORI	11
3.1 Microservices	11
3.2 Container	12
3.3 Kubernetes	13
3.2.1 Pod Object	14
3.2.2 Deployments Object	15
3.2.3 Nodes	15
3.2.4 Container Resource Request dan Limit	16
3.2.5 Metric Server	17
3.3. Autoscaling	17
3.3.1 Horizontal Pod Autoscaling	17
3.3.2 Vertical Pod Autoscaling	18

3.4.	Bi-LSTM.....	19
3.5.	Graph Neural Network.....	21
3.6.	Metode Evaluasi Forecasting	22
3.6.1	Mean Squared Error	22
3.6.2	Mean Absolute Error.....	22
3.6.3	Mean Absolute Percentage Error	23
3.6.4	R ² Score.....	23
3.7.	Metode Evaluasi Performa Autoscaling.....	24
3.7.1.	Service Level Objective	24
3.7.2.	Response Time	24
3.7.3.	Biaya operasional	25
3.7.4.	Skor Efisiensi Waktu Autoscaling	25
BAB IV METODOLOGI PENELITIAN		27
4.1	Tahapan Penelitian	27
4.2	Analisis Kebutuhan Sistem	29
4.2.1	Kebutuhan Perangkat Keras	29
4.2.2	Kebutuhan Perangkat Lunak	29
4.3	Perancangan Sistem	30
4.3.1	Persiapan Virtual Machine.....	31
4.3.2	Perancangan Klaster Kubernetes.....	32
4.3.3	Instalasi Service ke Klaster Kubernetes.....	33
4.3.4	Perancangan Adaptif Horizontal Autoscaling.....	33
4.3.4.1	Ingress	34
4.3.4.2	Microservices	34
4.3.4.3	Monitoring	35
4.3.4.4	Planner: Graphen HPA Operator	35
4.4	Pemodelan Machine Learning pada Graphen HPA	35

4.4.1	Pembuatan Dataset	36
4.4.2	Pembagian Dataset	37
4.4.3	Preprocessing	38
4.4.4	Tuning Lag Sliding Window	38
4.4.5	Tuning Hyperparameter	40
4.4.6	Pengujian dengan Skema Cross Validation	42
4.4.7	Forecasting	44
4.4.7.1	Prediksi Workload dengan Model Bi-LSTM	45
4.4.7.2	Prediksi Workload dengan Bi-LSTM-GNN	47
4.4.7.3	Integrasi Model	54
4.5	Implementasi dan Pengujian Sistem	55
4.5.1	Evaluasi Model Bi-LSTM dan Bi-LSTM-GNN	55
4.5.2	Evaluasi Performa Autoscaling	56
4.6	Analisis Data Hasil Pengujian	57
BAB V IMPLEMENTASI		58
5.1	Perangkat Sistem	58
5.1.1	Perangkat Keras	58
5.1.2	Perangkat Lunak	58
5.2	Implementasi Sistem	59
5.2.1	Provisioning Virtual Machine Melalui Google Kubernetes Engine	59
5.2.2	Konfigurasi Jaringan	60
5.2.3	Instalasi Istio	62
5.2.4	Instalasi Microservices	63
5.2.5	Instalasi Monitoring Service	65
5.2.6	Load Generator	67
5.2.7	Metrics Collector	70
5.2.7.1	Penggunaan CPU Microservices	70
5.2.7.2	Penggunaan Memory Microservices	70

5.2.7.3	Latency Microservices	71
5.2.7.4	Request Per Second (RPS) Microservices	72
5.2.7.5	Network Transmit Microservices	72
5.2.7.6	Network Receive Microservices	73
5.2.7.7	Mengintegrasikan Semua Metrics	74
5.2.8	Implementasi Pembagian Dataset Training/Validasi/Test	75
5.2.9	Implementasi Sliding Window	77
5.2.10	Implementasi Preprocessing	77
5.2.11	Implementasi Model Bi-LSTM dan Bi-LSTM-GNN	80
5.2.12	Implementasi Pelatihan Model	82
5.2.13	Implementasi Pengujian Model	84
5.2.14	Implementasi Pengujian Variasi Lag	86
5.2.15	Implementasi Tuning Hyperparamater	86
5.2.16	Implementasi Pengujian KFold	88
5.2.17	Implementasi Integrasi Model dengan Autoscaler Kubernetes	89
BAB VI HASIL DAN PEMBAHASAN		94
6.1	Hasil Dataset	94
6.1.2	Hasil Pembagian Dataset Training, Validasi dan Testing	95
6.2	Hasil Pengujian Sliding Window	96
6.3	Hasil Pengujian Hyperparameter Bi-LSTM	100
6.4	Hasil Pengujian Hyperparameter Bi-LSTM-GCN	101
6.5	Hasil Pengujian dengan Skema Cross-Validation (KFold)	103
6.6	Pengujian Autoscaling	107
6.6.1	Frontend Service	107
6.6.2	Checkout Service	109
6.6.3	Recommendation Service	109
6.6.4	Ads Service	110
6.6.5	Payment Service	111
6.6.6	Shipping Service	112
6.6.7	Email Service	114

6.6.8	Currency Service.....	115
6.6.9	Product Catalog Service.....	116
6.6.10	Cart Service.....	117
6.6.11	Biaya Operasional Cloud	119
6.7	Efisiensi Waktu Autoscaling.....	120
6.8	SLO Penggunaan CPU	121
6.9	SLO Response Time	122
BAB VII PENUTUP		125
7.1	Kesimpulan	125
7.2	Saran.....	125
DAFTAR PUSTAKA		127