

Big Data & Predictive Analysis Lanjut

Informasi & kontrak perkuliahan

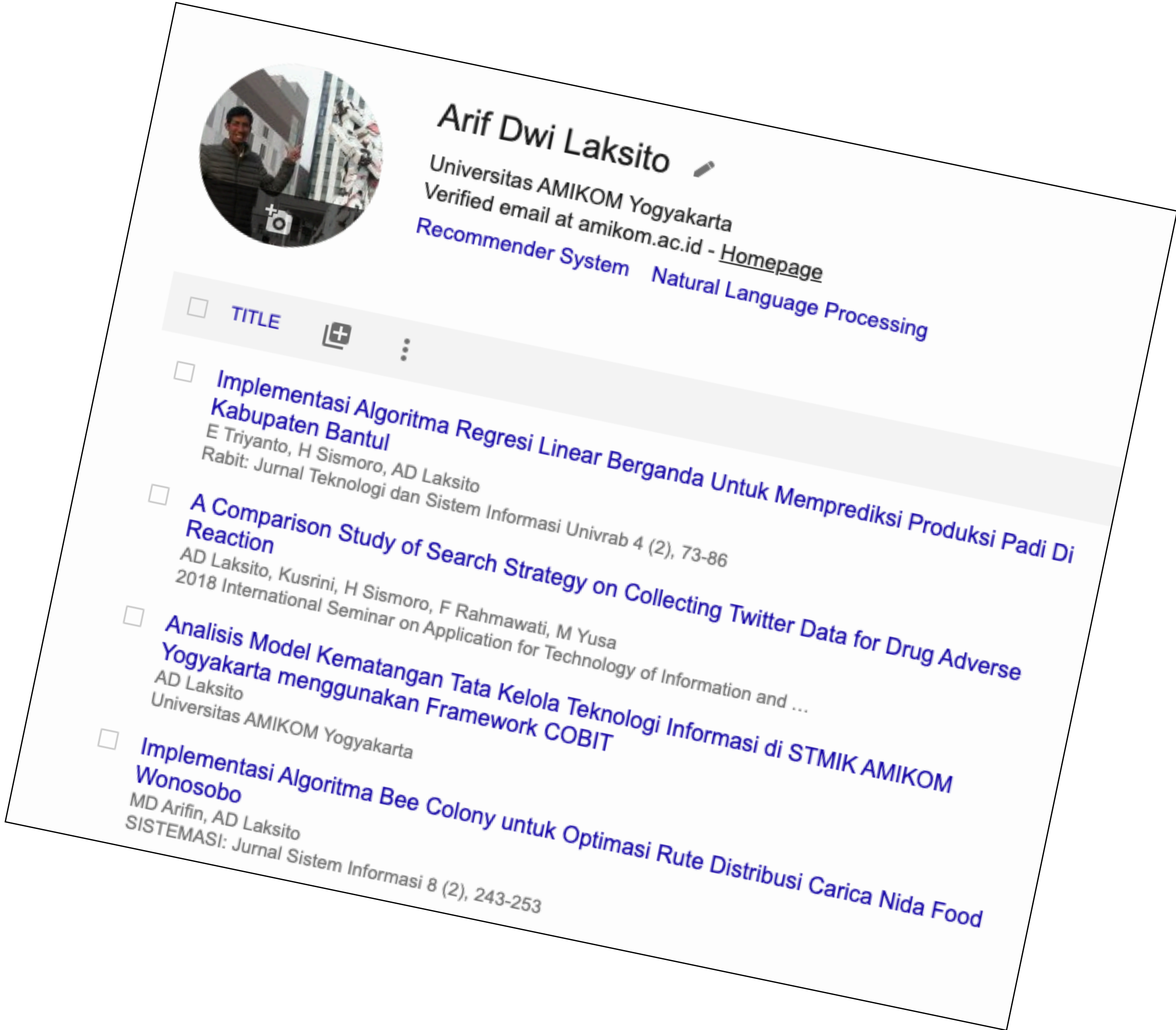
Arif Laksito, M. Kom


Arif Laksito, M. Kom



- Dosen tetap Prodi Informatika
- 2013, Master of Informatics Technology Universitas Amikom Yogyakarta
- 2006, Bachelor of Computer Science Universitas Gadjah Mada Yogyakarta
- Bidang studi: ML, NLP, RecSys
- Divisi Riset Direktorat Innovation Center Universitas Amikom
- ariflaksito.net

Arif Laksito, M. Kom








ariflaksito
Man Jadda Wajada

[f](#)[t](#)[in](#)


Join Research



TEAMWORK



EXCHANGE



INSPIRATION

COLLABORATION


SHARE


ASSIST


TRUST


SUPPORT

<https://ec.europa.eu/research-and-innovation/en/projects/success-stories/all/scaling-hurdles-global-science-collaboration>

For students who are interested in doing research in the **Recommender system (RS)** areas with me, there are several steps:

1. Read articles I wrote here:
<https://blog.ariflaksito.net/search/label/Recommendation%20system>
2. Search the newest papers related to the recommender system (it is about 10 – 20 papers) – you can find **RS handbooks** here
3. Discuss your ideas and send those to: arif.laksito@amikom.ac.id
4. Join the "recommender system group" at telegram:
https://t.me/joinchat/CH_iboMW7XdjZTA1

Prospective Students:

1. Basic knowledge in Programming (Python would be better)
2. Expected to study hard and learn new knowledge
3. Web/Mobile Development (optional)
4. Minimum GPA: 3.00

Available research topics:

1. Image extraction for content-based filtering RS
2. Music extraction recommender system
3. Model-based recommender system in several domains
4. Knowledge-based recommender system
5. Cross-domain recommender system
6. Deep learning in recommender system
7. Aux. information & latent factors in collaborative filtering RS

Other topics:

1. Text classification/Sentiment Analysis: RNN, LSTM, Attention, Transformer
2. Word embedding in NLP: Glove, Word2Vec, fastText, BERT
3. Machine learning model improvement using several datasets.
4. Neural Machine Translation

Benefits:

1. Free access datacamp & DQLab e-learning platform to improve your coding skills in R & Python related to Machine learning & Data Mining.
2. Intense discussion (twice a week)
3. Publication support (reputed journal or conference)

Aturan Perkuliahan

- 2 SKS Praktikum
- 1 x 90 menit setiap Minggu
- Maksimal terlambat 15 menit
- Minimal kehadiran 80% dari 14 x pertemuan
- Aktif dalam diskusi di perkuliahan
- Join Classroom dg kode:
 - #7 >> fs262v5
 - #8 >> fxasnlt

Deskripsi Matakuliah

Tujuan dari Matakuliah ini adalah mahasiswa mampu melakukan pengolahan data dalam ukuran besar menggunakan konsep komputasi terdistribusi menggunakan framework PySpark, serta mampu melakukan analisis data untuk menentukan model klasifikasi, regresi atau clustering yang sesuai

Materi pembelajaran

Sebelum UTS

- Introduction to PySpark
- Data Frame
- Manipulating data
- PySpark RDD

Setelah UTS

- PySpark MLlib
- Collaborative Filtering
- Classification
- Clustering

Capaian pembelajaran

1. Mahasiswa dapat mengimplementasikan **tools** pengolahan Big Data
2. Mahasiswa memahami penggunaan **Framework PySpark**
3. Mahasiswa mampu melakukan analisa data untuk menentukan **model Machine learning** pada Big Data

Spesifikasi kebutuhan

Menggunakan komputer/laptop dengan minimal spesifikasi:

- RAM minimal 2 GB (disarankan 4 GB)
- Operating System Windows 7, 8, 10, MacOS X atau Linux
- Konektivitas Internet baik dan stabil
- Memiliki akun Google Colab/Kaggle

Perangkat lunak

Selama pembelajaran membutuhkan beberapa perangkat lunak, seperti:

- Browser (disarankan Google Chrome)
- Python versi 3.x.x
- Framework PySpark

Komponen penilaian

- Presensi/Keaktifan/Quiz = 5%
- Tugas Mandiri = 15%
- Responsi 2x = 60%
- UTS + UAS = 20%

Klasifikasi Nilai:

$x > 80$: A

$\geq 60\% \ \& \ < 80$: B

$\geq 40\% \ \& \ < 60$: C

$\geq 20\% \ \& \ < 40$: D

< 20 : E