ABSTRAK

Efektivitas serta cakupan generalisasi metode deteksi anomali yang umum digunakan untuk data akuntansi berskala besar telah lama menjadi permasalahan yang diperdebatkan. Umumnya, performa pendekatan-pendekatan yang umum digunakan saat ini seringkali dinilai kurang memuaskan. Pendekatan-pendekatan yang lazim digunakan umumnya didasarkan pada kasus-kasus fraud lampau yang tidak dapat digeneralisasi dengan baik pada data-data baru. Terlebih lagi, permasalahan ini juga memiliki konsekuensi bisnis yang bersifat negatif apabila indikator temuan yang tidak tepat menyetir arah prosedur-prosedur lanjutan. Di saat yang bersamaan, perkembangan konsep serta penerapan teknologi kecerdasan buatan (AI) dan pembelajaran mesin (ML) sebagaimana telah didemonstrasikan pada bidang-bidang keilmuan lainnya sebaliknya merepresentasikan potensi solutif menjanjikan dalam menyelesaikan polemik tersebut. Penelitian ini dilakukan untuk menguji efektivitas pendekatan berbasis kecerdasan buatan/pembelajaran mesin yakni jaringan syaraf tiruan autoencoder dalam konteks deteksi anomali, yang juga dibandingkan dengan salah satu metode deteksi anomali akuntansi yang populer digunakan yaitu analisis digit hukum Benford, serta mengajukan suatu pendekatan baru yang menggabungkan kedua metode tersebut melalui suatu mekanisme multiplier yang didapatkan secara heuristik melalui eksperimen berulang. Dengan melakukan pengujian penerapan ketiga metode pada data nyata yang terdiri dari 500,000+ baris data jurnal akuntansi yang diperoleh dari tabel BKPF dan BSEG dari sistem ERP SAP suatu entitas nyata yang telah disamarkan dan disuntik dengan anomali sintetis, ditemukan bahwa pendekatan berbasis jaringan syaraf tiruan autoencoder sebagai metode yang paling efektif diukur dari sensitivitas (recall) serta dalam menyeimbangkan tradeoff precision dan recall melalui metrik F1-score. Temuan ini menggarisbawahi potensi menjanjikan penerapan pendekatan berbasis kecerdasan buatan/pembelajaran mesin terutama sekali pada konteks pengauditan internal. Pendekatan baru yang diajukan mendemonstrasikan wanprestasi dalam tujuannya untuk memperbaiki masalah recall dari model baseline autoencoder, kendatipun demikian perbedaan kinerja di antara keduanya tidak cukup signifikan untuk memberikan kesimpulan yang dapat digeneralisasi secara luas. Temuan-temuan ini menunjukkan tingginya potensi kinerja model autoencoder dalam implementasi yang mengutamakan performa recall dibandingkan dengan precision (misalnya dalam audit internal), sedangkan kondisi sebaliknya dapat dengan lebih baik diakomodasi menggunakan analisis digit pertama Hukum Benford (misalnya dalam audit eksternal).

Keywords: Deteksi Anomali . Akuntansi . Pengauditan . Autoencoder . Hukum Benford . Jaringan Syaraf Tiruan . Pembelajaran Mesin . Deep learning . Kecerdasan Buatan