KATALOG BPS: 1202031
ISSN: 2086–4132

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

TAHUN 6, VOLUME 2, DESEMBER 2014

UNIT PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT
SEKOLAH TINGGI ILMU STATISTIK
(UPPM-STIS)

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK
Journal of Statistical Application & Statistical Computing

No Publikasi / *Publication Number*: 02700.1004
Katalog BPS / *BPS Catalogue*: 1202031
No ISSN / *ISSN Number*: 2086-4132

Ukuran Buku / *Book Size*: 14,8 cm x 21,5 cm
Jumlah Halaman / *Number of Pages*: 139 + v

Diterbitkan oleh / *Published by*:
Sekolah Tinggi Ilmu Statistik
STIS-*Statistics Institute*

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

Alamat Redaksi : Sekolah Tinggi Ilmu Statistik
Jl. Otto Iskandardinata 64C
Jakarta Timur 13330
Telp. 021-8191437

TAHUN 6, VOLUME 2, DESEMBER 2014

KATALOG BPS: 1202031
ISSN: 2086-4132

JURNAL APLIKASI STATISTIKA & KOMPUTASI STATISTIK

FRAMEWORK UNTUK MENDETEKSI PEMALSUAN DATA PADA *MOBILE SURVEY*

Ibnu Santoso
Dosen Sekolah Tinggi Ilmu Statistik

Abstract

Interviewer falsifications are relevant problem faced by institutions conducting census and surveys around the world, including BPS-Statistics Indonesia. Falsified data may cause serious impact to generated statistics even though the proportion of falsified data is very small. Usage of Computer Assisted Personal Interviewing (CAPI) in field data collection has proven to improve efficiency and effectiveness. In addition, the use of CAPI is believed to be able to detect data falsification better. This is because CAPI devices can produce a variety of metadata that can not be obtained when using paper questionnaires. This study discusses relevant features to detect interviewer falsification in CAPI-based surveys, validates them, and uses them to identify interviewer falsification automatically using data mining techniques so that human supervisors can take further actions. After analyzing relevant features and conducting experiment, the result showed that unsupervised classification algorithm using simple 2-means clustering could have up to 70,5% accuracy, while supervised classification using logistic regression could have up to 88,5% accuracy.

Keywords: CAPI, interviewer falsification, unsupervised classification, supervised classification

I. PENDAHULUAN

Berdasarkan Undang-undang Nomor 16 Tahun 1997 tentang Statistik, salah satu peranan yang harus dijalankan oleh Badan Pusat Statistik (BPS) adalah menyediakan kebutuhan data bagi pemerintah dan masyarakat. Data ini banyak didapatkan dari sensus dan survei yang dilakukan oleh BPS. Dalam melaksanakan sensus dan survei tersebut, terutama jika skala kegiatannya cukup besar, BPS hampir selalu melibatkan mitra (tenaga *outsource*) yang direkrut oleh Koordinator Statistik Kecamatan (KSK) yaitu pegawai BPS yang bertanggung jawab di level kecamatan.

Visi BPS adalah sebagai pelopor data statistik terpercaya untuk semua. Agar data statistik dapat dipercaya oleh semua kalangan, kualitas data adalah suatu hal yang harus terjamin. Kualitas data hasil sensus dan survei ditentukan oleh banyak faktor. Diantara faktor-faktor tersebut adalah faktor responden dan pencacah. Kualitas data yang baik salah satunya dapat diharapkan dari kombinasi responden yang kooperatif memberikan jawaban yang akurat apa adanya dan dari petugas pencacah yang menguasai konsep dan definisi, memiliki etika dan teknik wawancara, serta memiliki kejujuran dan integritas yang tinggi.

Dalam prakteknya selalu ada hambatan dalam mencapai kualitas data yang diharapkan. Hambatan tersebut diantaranya berasal dari faktor responden yang sulit ditemui, responden yang tidak kooperatif dan memberikan jawaban sekenanya, dan responden yang tidak bisa memberikan jawaban dengan akurat. Sementara dari sisi petugas pencacah diantaranya petugas yang kurang menguasai etika dan teknik wawancara, kurang menguasai konsep dan definisi, dan petugas yang bekerja tidak sesuai prosedur operasional seperti melakukan kecurangan dengan mengisi sendiri seluruh atau sebagian isian kuesioner tanpa melakukan wawancara tatap muka dengan responden (Harrison, 1947). Kecurangan petugas dalam hal ini disebut juga sebagai pemalsuan data.

Pemalsuan data merupakan permasalahan relevan yang dihadapi oleh BPS. Dalam mengatasi permasalahan ini, BPS telah menjalankan program pengawasan dan monitoring berjenjang di setiap kegiatan survei dan sensus. Pengawasan berjenjang ini bertujuan untuk memastikan bahwa petugas bekerja sesuai dengan *Standard Operation Procedure* (SOP) sehingga data yang dihasilkan dapat dipertanggungjawabkan. Sebagai contoh pada kegiatan pendataan Sensus Penduduk 2010, satu tim pencacah terdiri dari seorang koordinator tim yang bertugas mengawasi tiga orang pencacah. Untuk mengawasi tim pencacah, masih ada petugas monitoring kualitas mulai dari level kabupaten/kota, provinsi, dan pusat.

Relevansi permasalahan pemalsuan data di BPS salah satunya dapat dilihat dari laporan resmi hasil monitoring kualitas Sensus Penduduk 2010 yang menyatakan bahwa ada indikasi pemalsuan data di beberapa daerah yang diawasi (BPS, 2010). Bukan hanya di BPS, permasalahan pemalsuan data juga merupakan kasus yang ditemui pada penyelenggaraan sensus dan survei oleh lembaga-lembaga statistik di berbagai negara, bahkan di negara maju. Berbagai studi dan laporan terkait tingkat pemalsuan data pada berbagai survei yang telah dilakukan dijabarkan pada Tabel 1.

Tabel 1 Berbagai Studi dan Laporan terkait Tingkat Pemalsuan data

No	Survei	Negara	Hasil Studi dan Laporan
1	*Monthly Current Population Survey* dan *Annual National Crime Survey*	Amerika	0.4% petugas memalsukan data. Padahal wawancara dilakukan oleh petugas professional (S. Bredl, P. Winker, dan K. Kotschau, 2008)
2	*Household Vacancy Survey*	Amerika	Tingkat pemalsuan data sebesar 6,5%. Petugas adalah staf temporer (S. Bredl, P. Winker, dan K. Kotschau, 2008)
3	*ALLBUS, German General Social Survey 1994*	Jerman	tingkat pemalsuan data sebesar 2,3% (A. Koch, 1995)
4	*Phone Survey*	Amerika	6% pewawancara mengakui telah memalsukan seluruh wawancara dan 13% pewawancara mengakui telah memalsukan sebagian data wawancara (P. Kiecker dan J. E. Nelson, 1996)
5	*US-National Health Interview Survey (NHIS)*	Amerika	3 dari 83 (3,6%) petugas yang dicurigai positif melakukan pemalsuan data (C. Hood dan M. Bushery, 1997)
6	*The 1997-98 Baltimore STD and Behavior Survey (BSBS)*	Amerika	7 dari 36 petugas (19,4%) melakukan pemalsuan data. 49% dari 451 wawancara yang dilakukan oleh 6 petugas diantaranya adalah palsu. 7 orang tersebut merupakan petugas yang belum memiliki pengalaman survei sebelumnya.(C. Turner dkk, 2002)
7	*American National Drug Survey on Drug Use and Health* (NSDUH)	Amerika	terdapat 3 orang petugas yang memalsukan data (J. Murphy dkk, 2004)
8	9 survey yang diselenggarakan *Census Bureau* antara 2005-2009	Amerika	143 dari 735 kasus wawancara yang dicurigai dipastikan palsu (C. Lawrence dan E. Love, 2010)

Di Indonesia belum ada studi khusus yang meneliti tingkat pemalsuan data suatu survei atau sensus maupun indikator-indikatornya. Penelitian tentang pemalsuan data menjadi sangat penting karena besarnya dampak yang dapat ditimbulkannya. Salah satu dampak dari pemalsuan data diantaranya dapat menyebabkan akibat yang serius untuk angka statistik yang dihasilkan dari data survei (S. Bredl, P. Winker, dan K. Kotschau, 2008). Untuk kegiatan survei, jika semakin banyak kasus pemalsuan data terjadi dan semakin besar perbedaan antara

nilai variabel yang sebenarnya dengan nilai variabel yang dipalsukan, maka akan semakin besar pula nilai bias dari hasil survei (J. Murphy dkk, 2004).

Efek pemalsuan data untuk statistik univariat mungkin tidak terlalu besar karena bagian yang dipalsukan tidak banyak dan data yang dipalsukan mungkin nilainya seperti data asli. Tetapi dari sebagian kecil data palsu tersebut sudah cukup untuk menghasilkan bias yang besar pada statistik multivariat (R. Schnell, 1991). Penelitian pada data German Socio Economic Panel (GSOEP) menemukan bahwa penyertaan data GSOEP palsu dalam regresi multivariat mengurangi efek training pada log upah kotor (*gross wages*) sebesar sekitar 80 persen, meskipun proporsi data palsu tersebut kurang dari 2,5 persen (J. Schrapler and G. Wagner, 2003). Oleh karena itu, untuk memenuhi salah satu persyaratan kualitas data maka tindakan pemalsuan data idealnya tidak boleh terjadi sama sekali dalam kegiatan pengumpulan data.

Bredl dkk. (2011) membagi studi tentang deteksi pemalsuan data menjadi dua jenis: (1) *Ex-ante study*, yaitu studi yang membahas tentang metode-metode deteksi yang diaplikasikan saat pelaksanaan lapangan dengan tujuan untuk mengidentifikasi petugas mana yang melakukan kecurangan. (2) *Ex-post study*, yaitu studi yang mengaplikasikan berbagai indikator pada suatu dataset dengan kasus-kasus kecurangan yang telah diketahui dengan tujuan mengidentifikasi indikator-indikator yang membedakan antara data yang diperoleh dengan jujur dan data yang telah dipalsukan.

Ex-ante study menjadi penting karena pelaksanaan lapangan suatu kegiatan sensus atau survei memerlukan suatu strategi proses monitoring atau pengawasan yang baik dan terencana untuk memastikan kegiatan lapangan berjalan sesuai dengan SOP. Proses monitoring saat kegiatan pelaksanaan pendataan lapangan sedang berlangsung juga berguna untuk mendeteksi dan meminimalisir kejadian pemalsuan data. M Rita Tisshen (2008) telah merangkum berbagai teknik/metode monitoring klasik beserta kelebihan dan kekurangannya sebagaimana dijelaskan pada Tabel 2. Namun, teknik klasik tersebut memiliki kekurangan dalam skala penerapannya, yaitu jumlah tenaga pengawas atau supervisor maupun dana yang ada dapat menjadi keterbatasan jika harus mengawasi seluruh petugas dan memantau seluruh kegiatan wawancara. Maka hampir semua prosedur pengawasan dengan teknik klasik ini hanya dilakukan secara sampel, tidak dapat menjangkau seluruh populasi.

Ex-post study memiliki arti penting ketika indikator-indikator yang berhubungan dengan pemalsuan data dapat diidentifikasi dan memiliki pengaruh yang kuat. Contoh dari hasil *ex-post study* seperti dijelaskan oleh Chrishelle Lawrence dan Elizabeth Love (2010) adalah petugas yang minim pengalaman/belum memiliki pengalaman pendataan lapangan sama sekali dan petugas yang berasal dari wilayah tertentu memiliki kecenderungan yang lebih tinggi dalam melakukan pemalsuan data. Indikator ini dapat menjadi input dalam melakukan

pengawasan/monitoring untuk survei berikutnya. Sebagai contoh, pada survei berikutnya, petugas yang memenuhi kriteria dalam indikator tersebut mendapatkan porsi supervisi/pengawasan yang lebih tinggi daripada petugas yang sudah berpengalaman.

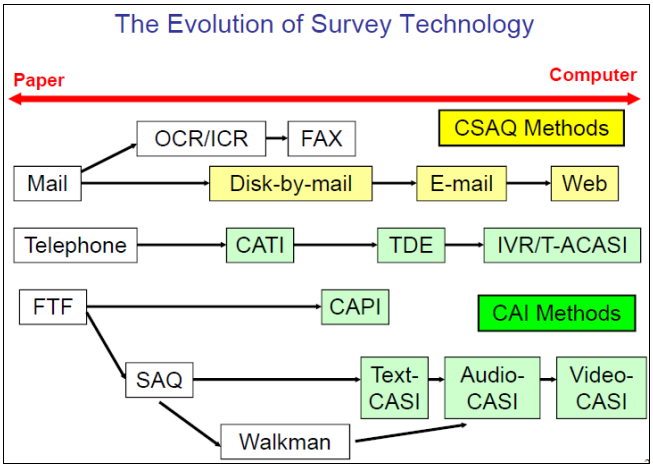
Tabel 2. Berbagai Teknik Klasik dalam Pengawasan Kegiatan Lapangan (M Rita Tisshen, 2008)

Teknik	**Kelebihan**	**Kekurangan**
Observasi Langsung	Memberikan gambaran mendetail	-Dapat mempengaruhi pewawancara dan responden -Mahal dari segi biaya
Diskusi dengan pewawancara setelah selesai wawancara	Murah dari segi biaya	Hanya mendapatkan sudut pandang pewawancara
Verifikasi responden lewat telepon atau wawancara ulang secara langsung	Mendapatkan sudut pandang responden	-Membebani responden -Mahal dari segi biaya
Melakukan review pada data respons dan waktu	Efektif untuk mengawasi kualitas data	Terbatasnya informasi tentang performa wawancara
Merekam pembicaraan dengan *tape recorder*	Informatif	-Dapat mengganggu -Canggung jika *tape* harus di stop dan start -Biaya dan usaha ekstra untuk perlengkapan dan pengiriman *tape*

Perkembangan teknologi merambah semua bidang tidak terkecuali pada teknologi pendataan lapangan sebagaimana diilustrasikan pada Gambar 1. Perkembangan teknologi memungkinkan wawancara tidak lagi dilakukan dengan tatap muka melainkan bisa lewat telepon dan e-mail. Kemudian media pendataan juga mengalami perubahan dari media kertas menjadi media komputer. Teknologi pengolahan juga berkembang dari yang tadinya isian kuesioner kertas di-input ke komputer oleh petugas pengolahan menjadi di-scan oleh mesin scanner. Teknologi terkini memungkinkan wawancara dilakukan menggunakan media komputer genggam kemudian hasil wawancara langsung dikirimkan ke server.

Meskipun dari segi biaya termasuk yang menghabiskan biaya paling besar dibanding moda lain, survei tatap muka tetap menjadi pilihan BPS karena memiliki tingkat respon yang paling baik. Survei tatap muka juga menjadi kebutuhan bagi responden tertentu, misalnya responden yang buta huruf, responden yang berdomisili di daerah terpencil serta responden

yang tidak dapat dijangkau surat maupun telepon. Dalam perkembangannya di BPS, survei tatap muka yang tadinya menggunakan media kertas (*Paper and Pencil Interviewing*, PAPI) secara perlahan berubah menggunakan media komputer (*Computer Assisted Personal Interview*, CAPI) yang ditandai dengan pembelian perangkat berbasis Android sebanyak 1300 unit khusus untuk pendataan lapangan.



Gambar 1. Perkembangan Teknologi Survei (University of Maryland, 2013)

Survei berbasis CAPI pertama yang dilakukan oleh BPS adalah Survei Penggunaan Tembakau Indonesia (SPTI) 2011 (BPS, 2011). Kemudian pada bulan Juni 2013 BPS melakukan Pendataan Pilot Survei Pengetahuan, Sikap, dan Perilaku Kesiapan Menghadapi Bencana di Kota Padang, Sumatera Barat. Dalam pelaksanaan kegiatan yang kedua ini, BPS menggunakan tablet berbasis Android untuk kegiatan lapangannya. Penggunaan tablet berbasis Android dalam pengumpulan data bertujuan untuk meminimalisir kesalahan yang menyebabkan data tidak bersih. Survei berbasis CAPI yang dilakukan oleh BPS belum dilengkapi dengan mekanisme untuk mendeteksi pemalsuan data secara cepat, yaitu mekanisme yang dapat mendeteksi pemalsuan data yang dilakukan oleh petugas pada saat kegiatan pengumpulan data masih berjalan.

Moda CAPI masih merupakan hal yang baru bagi BPS. Selain memberikan keuntungan lebih dalam proses pengolahan data, pengumpulan data dengan moda CAPI dipercaya dapat memberikan keuntungan lebih dalam proses monitoring dan pendeteksian pemalsuan data oleh petugas. Hal ini karena dengan moda CAPI dapat diperoleh metadata yang menjelaskan kapan, dimana, dan bagaimana data respon diperoleh.

Pemalsuan data dapat terjadi pada kegiatan pengumpulan data, terlepas dari apapun moda pengumpulan data yang digunakan. Untuk itu diperlukan suatu strategi pengawasan lapangan yang baik karena keterbatasan jumlah dan kapabilitas supervisor manusia dalam

melakukan pengawasan. Dengan menggunakan moda CAPI, supervisor manusia dapat menjadi lebih terbantu karena pengawasan atau monitoring dapat dilakukan secara sistemik, lebih terfokus, dan menjangkau seluruh populasi, tidak lagi secara sampel seperti yang ditemukan pada keterbatasan teknik klasik dalam melakukan pengawasan sebagaimana dijelaskan pada Tabel 2 di atas.

Dalam penelitian ini pertama penulis berusaha untuk mengkaji fitur-fitur apa saja yang dapat digunakan untuk mendeteksi pemalsuan data. Kemudian setelah mendapatkan fitur-fitur tersebut, penulis mencoba menganalisis kebutuhan sistem pendeteksi pemalsuan data dan membuat framework untuk sistem tersebut. Penulis melakukan eksperimen dengan menerapkannya pada suatu survei nyata yang dilakukan oleh BPS yang melibatkan petugas yang berpengalaman dan mengukur apakah sistem pendeteksi pemalsuan wawancara dapat berjalan dengan baik.

Berdasarkan uraian pada latar belakang di atas, maka masalah pada penelitian dapat dirumuskan sebagai berikut:

1. Apa saja fitur yang relevan dan dapat dijadikan rujukan untuk mendeteksi pemalsuan data?
2. Bagaimana sistem yang dapat mendeteksi kecurangan petugas dapat diterapkan pada survei berbasis CAPI?

Tujuan penelitian ini adalah:

1. Melakukan identifikasi fitur yang relevan untuk mendeteksi pemalsuan data oleh petugas pencacah.
2. Merancang dan membangun sistem *prototype* pendeteksi pemalsuan data untuk diterapkan pada instrumen *tablet phone,* melakukan ujicoba, dan mengevaluasi hasil ujicoba.

Ruang lingkup penelitian ini adalah:

1. Hanya melakukan pendeteksian pemalsuan data secara otomatis saja, tidak membahas pada pencegahan sebelum terjadi dan tindakan setelah terjadi.
2. Tidak membahas tindakan kecurangan yang dilakukan oleh selain petugas pewawancara.
3. *Prototype* dibangun pada perangkat *tablet phone* berbasis Android.

II. TINJAUAN PUSTAKA

Pemalsuan data

AAPOR (2003) mendefinisikan pemalsuan data (*interviewer falsification*) sebagai perbuatan sengaja dari petugas untuk tidak mematuhi petunjuk dan instruksi pencacahan yang tidak dilaporkan oleh petugas itu sendiri dan dapat berakibat pada kontaminasi data. “Sengaja” berarti bahwa petugas melakukan secara sadar perbuatan menyimpang tersebut.

Pemalsuan data yang dimaksud diantaranya:

a. memalsukan seluruh atau sebagian wawancara – dengan cara mengisi sendiri isian kuesioner dan melaporkannya sebagai jawaban responden;
b. sengaja salah melaporkan kode disposisi dan memalsukan data proses (misalnya, pencatatan kasus *non response*, melaporkan upaya menghubungi responden secara fiktif);
c. sengaja menulis kode jawaban yang berbeda dari responden untuk menghindari pertanyaan berikutnya;
d. sengaja melakukan wawancara pada rumah tangga yang bukan sampel, untuk mengurangi usaha yang diperlukan dalam menyelesaikan proses wawancara;
e. sengaja membuat laporan palsu tentang proses pengumpulan data kepada manajemen survei.

Kecurangan yang dimaksud tidak termasuk kesalahan umum dan tidak disengaja, misalnya kesalahan pengukuran dalam situasi tanya jawab atau kesalahan oleh pewawancara dalam merekam jawaban responden karena untuk tidak memahami atau ingat protokol wawancara. Dengan demikian, menentukan bahwa tindakan kecurangan telah terjadi melibatkan beberapa penilaian tentang niat dari petugas itu sendiri.

Ada beberapa alasan mengapa petugas pencacah melakukan pemalsuan data. Diantaranya disebabkan oleh kuesioner yang panjang, pertanyaan yang kompleks, responden yang sulit dijangkau, dan faktor eksternal seperti cuaca dan kondisi masyarakat (L. Crespi, 1945). Biasanya petugas pencacah tidak memiliki kepedulian kuat untuk mendapatkan kualitas data yang tinggi. Petugas tidak terlibat dalam perencanaan survei atau pengembangan kuesioner dan jarang sekali ada petugas yang menguasai etika riset ilmiah (AAPOR, 2003). Petugas harus melakukan wawancara dengan responden yang tidak dikenal untuk mendapatkan informasi personal yang mungkin bersifat sensitif sehingga kemungkinan akan muncul rasa sungkan. Petugas mungkin juga dihadapkan pada skema pemberian honor yang hanya berdasarkan pada jumlah responden yang diwawancarai (A. Kennickell, 2002), suatu skema yang bisa menyebabkan pemahaman bahwa kuantitas wawancara lebih berarti dibandingkan kualitasnya (A. Bennet, 1948).

Ciri-ciri Data Palsu

Data-data yang dipalsukan memiliki sifat-sifat tertentu. Berdasarkan studi pustaka, Benjamin B. (2012) telah menghimpun 15 ciri-ciri data palsu dari berbagai makalah. Kemudian ciri-ciri data yang dipalsukan tersebut dibagi dalam empat kategori karakteristik yaitu:

1. Menurut karakteristik bahwa petugas ingin usaha minimal dalam menyelesaikan wawancara:
   a. *Fast interview, w*awancara dilakukan dalam waktu yang cepat

b. *Many missing unit,* banyak terdapat jawaban yang kosong

c. *Interview surge,* terjadi lonjakan jumlah wawancara yang selesai dalam waktu yang singkat

d. *Close to deadline,* mendekati batas akhir penyelesaian wawancara

e. *Short paths through survey,* petugas memilih "rute" pertanyaan yang pendek dalam kuesioner

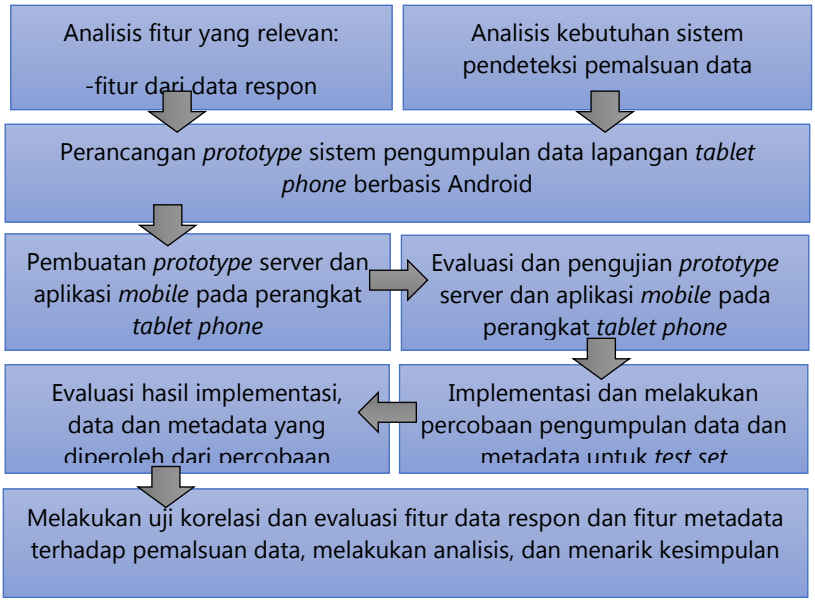
f. *Unusual time of day,* wawancara dilakukan pada waktu yang tidak biasa, terlalu pagi atau terlalu petang

g. *Many incomplete interview,* banyak wawancara yang tidak selesai

2. Menurut karakteristik bahwa petugas ingin menghindari pengawasan:
   a. *Missing telephone number,* Nomor telepon responden pada kuesioner sengaja dikosongkan
   b. *Low data variance,* variasi data rendah
   c. *Few missing units,* hanya sedikit jawaban yang kosong
3. Menurut karakteristik bahwa petugas tidak mengetahui sebaran distribusi populasi data yang sebenarnya:
   a. *Bad fit to benford's law,* kurang sesuai dengan hukum distribusi angka benford
   b. *Unusual data,* adanya data yang tidak biasa
   c. *Rare response combination,* kombinasi jawaban yang langka
4. Menurut karakteristik lain:
   a. *Low time variance,* variasi waktu penyelesaian wawancara kecil
   b. *Long interview,* wawancara dilakukan dalam waktu yang lama

III. METODE PENELITIAN

Metodologi yang digunakan dalam penelitian ini menggunakan *Design Science Research Methodology* (DSRM) *for Information System Research.* Alur dan kerangka pemikiran yang digunakan dalam penelitian ini dapat dilihat pada Gambar 2. Penelitian dimulai dengan melakukan analisis fitur yang relevan melalui berbagai studi pustaka dan analisis kebutuhan sistem yang dapat mendeteksi pemalsuan data secara otomatis dan diakhiri dengan melakukan uji korelasi dan evaluasi fitur data respon dan fitur metadata terhadap pemalsuan data, melakukan analisis, dan menarik kesimpulan untuk mencoba menjawab pertanyaan penelitian.



Gambar 2. Alur Pelaksanaaan Penelitian

IV. HASIL DAN PEMBAHASAN

Berikut adalah berbagai fitur metadata yang dapat digunakan untuk mendeteksi data palsu yang dirangkum dari berbagai literatur.

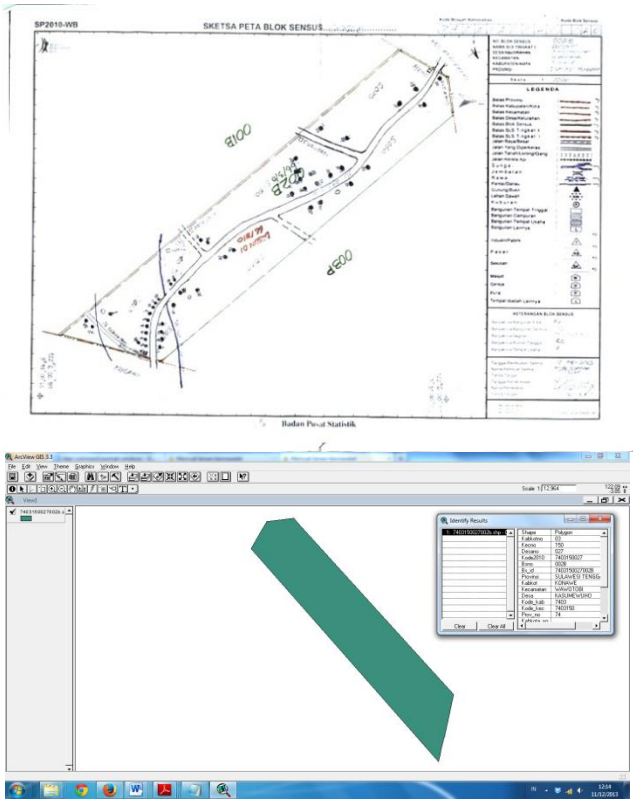
Fitur-fitur yang Relevan

1. **Lokasi**

Secara teoritis, wawancara yang jujur dilakukan di tempat atau lokasi yang benar. Pemalsuan data bisa dilakukan di tempat yang benar bisa juga tidak. Jadi, lokasi atau tempat dapat memegang peranan yang penting dalam mendeteksi pemalsuan data. Perangkat tablet phone saat ini sudah dilengkapi dengan *Global Positioning System* (GPS) untuk mendeteksi lokasi dimana perangkat tersebut berada. Lokasi yang didapat dari GPS dinyatakan dalam koordinat *latitude* dan *longitude*.

Pendeteksian lokasi sangat berguna jika terdapat data lokasi koordinat responden. BPS tidak mempunyai data lokasi koordinat responden, tetapi BPS memiliki peta wilayah kerja yang sudah memiliki lokasi geografis. Unit terkecil peta wilayah kerja adalah peta blok sensus. Dalam satu blok sensus terdiri dari 80-100 rumah tangga. Peta blok sensus merupakan perlengkapan standar yang harus dibawa oleh petugas pada saat melakukan pengumpulan data (Gambar 3). Penggunaan peta memastikan bahwa data diambil dari tempat yang benar. Peta blok sensus berbentuk polygon dan dapat berbentuk peta digital. peta digital yang dimiliki BPS disimpan dalam bentuk *shapefile*.

Pendekatan pemeriksaan lokasi bisa dipastikan dengan melihat apakah suatu titik koordinat perangkat *tablet phone* pada saat melakukan wawancara berada dalam *polygon* peta blok sensus atau tidak. Jika titik koordinat tidak berada dalam *polygon*, bahkan jauh, maka kebenaran wawancara dapat dicurigai. Selain melihat titik dalam *polygon*, dapat juga dilihat pergerakan dari perangkat *tablet phone* tersebut. Secara teknis, pendeteksian satu titik koordinat berada dalam *polygon shapefile* dapat dilakukan menggunakan *library osgeo* dan *iphyton*. Dengan input koordinat lokasi *latitude* dan *longitude*, output program menghasilkan titik koordinat tersebut berada di *polygon* yang mana. Bila output *polygon* sama dengan wilayah kerja, maka wawancara dilakukan di wilayah kerja yang benar.



Gambar 3. Peta Analog dan Digital BPS

2. Date Stamps

Date Stamps adalah tanggal kapan wawancara dilakukan. Pada pendataan BPS yang menggunakan kertas sebagaimana ditunjukkan pada Gambar 4, biasanya terdapat kolom kapan wawancara dilakukan dan kapan pemeriksaan oleh pengawas dilakukan. *Date Stamps* sangat berguna dan dapat memberikan informasi dalam satu hari berapa kuesioner yang bisa diselesaikan oleh petugas. Jika terlalu banyak wawancara yang diselesaikan dalam satu hari melebihi batas normal, maka petugas dapat dicurigai melakukan pemalsuan data. Indikasi

pemalsuan data juga dapat dinilai ketika banyak wawancara yang diselesaikan mendekati deadline pendataan lapangan.

III. KETERANGAN PETUGAS

Uraian	Pencacah	Pengawas/Pemeriksa
1. Nama	RAHMIN	MITA
2. Jabatan	1. Staf BPS Provinsi ③KSK 2. Staf BPS Kab/Kota 4. Mitra [3]	1. Staf BPS Provinsi 3. KSK 2. Staf BPS Kab/Kota 4. Mitra [2]
3. Tanggal Pencacahan/Pemeriksaan	Tanggal [0][4] Bulan [0][7]	Tanggal [1][7] Bulan [0][7]
4. Tanda Tangan		

Gambar 4. Contoh Date Stamps

Kelemahan penggunaan kuesioner kertas, data *date stamp* ini dapat dipalsukan dengan mudah. Petugas dapat mengisi tanggal dilakukan wawancara sesuka hati. Mekanisme kontrol ada pada pengawas yang melakukan pengawasan petugas pencacah. Jika menggunakan perangkat *tablet phone*, *date stamp* dapat dikirimkan bersamaan dengan pengiriman data. Meski dapat juga dipalsukan, namun pengubahan *date stamp* pada perangkat *tablet phone* lokal, *date stamp* yang akurat dapat diperoleh menggunakan *time server internet*.

3. Time Stamps

Time Stamps menginformasikan pada jam berapa suatu wawancara dilakukan. Indikasi pemalsuan data jika beberapa wawancara memiliki interval *time tamps* yang terlalu rapat, atau wawancara terjadi pada waktu yang tidak lazim (misalnya antara jam 12 malam hingga jam 6 pagi).

Untuk mendapatkan metadata *time stamp* yang lebih akurat dapat digunakan *tablet phone* sebagai instrumen pendataan. Sama seperti *date stamps*, informasi *time stamp* juga bisa didapat dari kuesioner kertas, tetapi *time stamp* tersebut dapat dengan mudah dipalsukan oleh petugas.

4. Duration

Duration data secara umum adalah metadata yang berkaitan dengan lamanya wawancara atau durasi wawancara. Secara umum, *duration data* menjelaskan berapa lama satu wawancara dilakukan. Dalam beberapa literatur, timing data dipandang sebagai pemrediksi yang kuat untuk mendeteksi pemalsuan wawancara. Wawancara yang baik biasanya berlangsung dalam durasi waktu yang normal. Indikasi pemalsuan data jika wawancara selesai terlalu cepat, kemungkinan jika petugas tidak memalsukan data maka kurang menggali jawaban responden sehingga data yang diperoleh menjadi kurang berkualitas. Jika wawancara selesai terlalu lama dari waktu normal, mungkin petugas yang bersangkutan membutuhkan

pelatihan tambahan yang mengajarkan bagaimana menggunakan waktu wawancara dengan efisien.

Duration data ini dapat dijabarkan lagi tidak sekedar durasi keseluruhan wawancara. Penggunaan *tablet phone* memungkinkan aplikasi merekam dari keseluruhan lama waktu wawancara tersebut, berapa lama durasi petugas menanyakan pertanyaan dan berapa lama durasi responden menjawab. Pertanyaan yang sulit juga dapat membantu memprediksi karena waktu yang dibutuhkan oleh responden akan lebih lama untuk menjawabnya.

5. Behavioral Data

Behavioral data menjelaskan bagaimana pola perilaku interaksi antara petugas dengan responden. Misalnya petugas yang banyak melakukan *swipe* terlalu banyak dalam waktu yang singkat mungkin hanya berinteraksi dengan perangkatnya, tidak dengan responden.

Behavioral data misalnya jumlah *swipe*, *click*, *next*, *prev*, melakukan editing jawaban, dan menekan tombol bantuan. Penelitian terkini tentang pemalsuan data oleh Birnbaum dkk (2013) menggunaan *behavioral data* untuk mendeteksi pemalsuan data. Dalam studi tersebut, penggunaan *supervised learning* terhadap *behavioral data* dapat memberikan akurasi yang tinggi meski petugas tahu bagaimana sistemnya bekerja dan diberikan insentif lebih jika mampu memalsukan data tanpa diketahui oleh server.

Setelah mempelajari berbagai literatur, Birnbaum (2012) merangkum 15 karakteristik dari data-data yang dipalsukan. Kemudian dari karakteristik-karakteristik tersebut dapat dipetakan dengan fitur-fitur untuk mendeteksinya. Hasil pemetaan tersebut dapat dilihat pada Tabel 3.

Tabel 3. Pemetaan Ciri Data Palsu dan Fitur untuk Mendeteksinya

No	Karakteristik	Fitur Metadata					Fitur Data			
		1	2	3	4	5	6	7	8	9
1	*Fast interview*	-	-	√	√	-	-	-	-	-
2	*Many missing unit *)*	-	-	-	-	-	-	-	-	-
3	*Interview surge*	-	√	-	-	-	-	-	-	-
4	*Close to deadline*	-	√	-	-	-	-	-	-	-
5	*Short paths through survey*	-	-	-	-	-	-	-	-	√
6	*Unusual time of day*	-	-	√	-	-	-	-	-	-
7	*Many incomplete interview*)*	-	-	-	-	-	-	-	-	-
8	*Missing telephone number*	-	-	-	-	√	-	-	-	-
9	*Low data variance*	-	-	-	-	-	-	√	-	-
10	*Few missing units *)*	-	-	-	-	-	-	-	-	-
11	*Bad fit to benford's law*	-	-	-	-	-	√	-	-	-
12	*Unusual data*	-	-	-	-	-	-	-	√	-
13	*Rare response combination*	-	-	-	-	-	-	-	√	-
14	*Low time variance*	-	-	-	√	-	-	-	-	-
15	*Long interview*	-	-	-	√	-	-	-	-	√

Ket: 1=Lokasi, 2=Date Stamp, 3=Time Stamp, 4=Duration, 5=Behavioral Data, 6=Benford Law, 7=Variability, 8=Data rarity, 9=Jumlah skip

*) Tidak seperti kuesioner kertas, pada CAPI meminimalisir kemungkinan wawancara yang tidak selesai dan tidak ada unit yang tidak terisi.

Dari hasil pemetaan di Tabel 3 terlihat bahwa dari 15 sifat data palsu, 3 diantaranya sudah dapat diatasi dengan penerapan CAPI. 7 sifat data palsu dapat dilihat menggunakan metadata dan sisanya menggunakan data respon. Secara keseluruhan sifat data palsu dapat dideteksi terutama cukup dengan menggunakan fitur *date stamp*, *time stamp*, *timing data*, dan analisis data respon sederhana. Fitur lain seperti lokasi, dan *behavioral data* sebagaimana disebutkan sebelumnya diyakini dapat membuat deteksi lebih akurat.

Implementasi

Uji coba sistem pendeteksi pemalsuan data ini dilakukan pada Survei Perilaku Peduli Lingkungan Hidup (SPPLH). SPPLH dipilih karena merupakan survei BPS yang waktu penyelesaiannya di responden relatif tidak terlalu lama.

Kuesioner survei SPPLH dibagi dalam 13 blok dan 36 nomor pertanyaan. Jumlah kolom yang dibutuhkan dalam database untuk menyimpan satu record sejumlah 138 kolom. Kuesioner SPPLH memiliki banyak pertanyaan bersyarat (*conditional*) yaitu pertanyaan yang hanya ditanyakan jika jawaban tertentu diberikan pada pertanyaan sebelumnya. Jumlah variabel bersyarat pada kuesioner SPPLH berjumlah sekitar 51 variabel dari 138 variabel.

Karena diterapkan pada lingkungan percobaan, beberapa fitur tidak dapat digunakan seperti fitur date stamp dan time stamp yang berfungsi mencatat tanggal dan waktu dilakukan wawancara. Oleh karena itu digunakan fitur lain yang masih relevan dengan lingkungan percobaan ini sebagaimana dijelaskan pada Tabel 4.

1. Pembuatan Aplikasi

Aplikasi dikembangkan untuk perangkat Samsung Galaxy Tab 2 P3100 dengan spesifikasi sebagai berikut:

- Layar 7.0 inch
- Prosesor 1GHz Dual Core
- Ram 1 GB
- OS Android 4.0.3 Ice Cream Sandwich

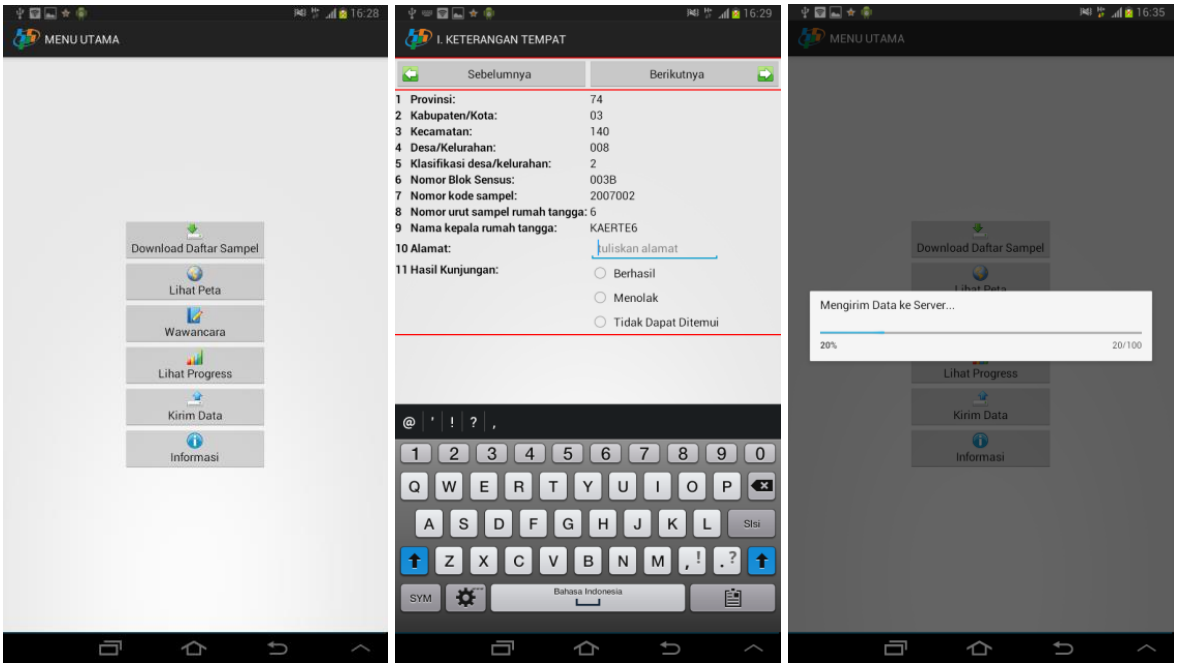
Aplikasi dilengkapi dengan fitur yang dapat merekam data dan metadata sebagaimana dibahas pada sub bab sebelumnya. Gambar 5 memperlihatkan *screenshot* dari aplikasi yang telah dibuat.

Aplikasi server dikembangkan menggunakan bahasa pemrograman PHP dengan fungsionalitas sebagai berikut:

1. Otentikasi login pengguna
2. Mentransfer daftar rumah tangga yang harus dicacah oleh petugas
3. Menampilkan kemajuan pelaksanaan pencacahan lapangan untuk setiap petugas
4. Menyimpan seluruh data respon dan metadata yang diunggah oleh client ke database

Tabel 4 Fitur yang Digunakan

No	Fitur	Penjelasan
1	Dalamblok	Apakah wawancara dilakukan di dalam blok sensus yang ditentukan atau tidak
2	Totalwaktu	Total waktu yang dibutuhkan untuk menyelesaikan wawancara
3	Totalwaktubertanya	Total waktu yang dibutuhkan untuk menanyakan pertanyaan
4	Reratawaktubertanya	Rata-rata waktu yang dibutuhkan untuk menanyakan pertanyaan (Totalwaktubertanya dibagi dengan jumlah layar pertanyaan yang dilalui)
5	Totalwaktujawab	Total waktu yang dibutuhkan untuk menjawab pertanyaan
6	Reratawaktujawab	Rata-rata waktu yang dibutuhkan untuk menjawab pertanyaan (Totalwaktujawab dibagi dengan jumlah layar pertanyaan yang dilalui)
7	Jumlahcontendit	Total berapa kali petugas mengubah-ubah jawaban dalam satu layar pertanyaan
8	Jumlahnoncontedit	Total berapa kali petugas tidak mengubah jawaban setelah layar muncul
9	Jumlahbersyarat	Total berapa kali layar pertanyaan bersyarat muncul
10	Waktubersyarat	Total waktu yang dihabiskan dalam layar pertanyaan bersyarat
11	Berikutnya	Total berapa kali tombol “berikutnya” ditekan
12	Sebelumnya	Total berapa kali tombol “sebelumnya” ditekan
13	Bantuan	Total berapa kali tombol “kalkulator” ditekan



Gambar 5. Aplikasi SPPLH

2. Pelatihan Petugas dan Pengumpulan Data

Pelatihan Petugas dilakukan selama satu hari kerja. Satu hari pelatihan dibagi menjadi dua sesi, sesi pertama pemahaman konsep dan definisi yang digunakan pada SPPLH sehingga petugas dapat bertanya dengan lebih spesifik dan efektif. Sesi kedua adalah pelatihan penggunaan perangkat tablet.

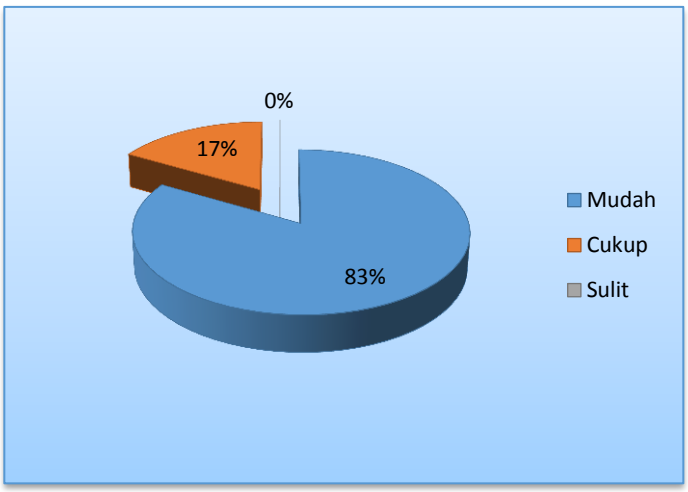
Peserta terdiri dari 30 orang, yang dibagi dalam 15 tim. Satu orang petugas mengumpulkan data 20 rumah tangga dengan pembagian 10 rumah tangga hasil pencacahan yang sebenarnya dan 10 rumah tangga hasil prediksi. 5 rumah tangga prediksi dilakukan pada saat pelatihan dan 5 sisanya dilakukan setelah mencacah 10 rumah tangga sebenarnya untuk mendapatkan pengetahuan dan pengalaman dalam mencacah rumah tangga yang sebenarnya. Sehingga total ada data 300 rumah tangga sebenarnya dan 300 rumah tangga prediksi. Metode eksperimen ini mirip dengan yang dilakukan Brinbaum dkk (2013) dalam studinya.

Hasil Tes Dan Evaluasi

1. Survei Penerimaan Pengguna

Setelah aplikasi diujicoba oleh pengguna, dilakukan survei penerimaan pengguna untuk melihat kemudahan penggunaan aplikasi. Gambar 6 menjelaskan kemudahan penggunaan aplikasi menurut pengguna dimana 25 orang (83%) menyatakan aplikasi mudah digunakan tanpa kendala dan sisanya 5 orang (17%) menyatakan bahwa

aplikasi cukup mudah digunakan. Tidak ada pengguna aplikasi yang mengeluhkan sulit menggunakan aplikasi.



Gambar 6. Kemudahan Penggunaan Aplikasi

2. Evaluasi Fitur

Dari 13 fitur yang digunakan di atas, kemudian dihitung korelasinya dengan pemalsuan data. Tabel 5 memperlihatkan nilai korelasi untuk masing-masing fitur.

Dari tabel 5 diperoleh data bahwa 7 fitur memiliki korelasi yang sedang hingga kuat. Sisanya tidak ada korelasi atau sangat lemah. Terlihat bahwa fitur yang memiliki korelasi yang kuat dan sedang kebanyakan berhubungan dengan durasi. Satu fitur *help* bahwa petugas yang memalsukan data ternyata tidak banyak menggunakan tombol bantuan.

Tabel 5. Nilai Korelasi Fitur (r) Terhadap Pemalsuan Data

Fitur	**r**	**Deskripsi**
Totalwaktu	0.539	Kuat
Reratawaktujawab	0.506	Kuat
Totalwaktujawab	0.499	Kuat
Totalwaktubertanya	0.489	Sedang
Reratawaktubertanya	0.470	Sedang
Bantuan	0.467	Sedang
Waktubersyarat	0.382	Sedang
Jumlahcontedit	-0.126	Lemah
Berikutnya	0.079	Tidak ada/sangat lemah
Sebelumnya	0.061	Tidak ada/sangat lemah
Jumlahnoncontedit	0.057	Tidak ada/sangat lemah
Jumlahbersyarat	0.055	Tidak ada/sangat lemah

3. Unsupervised Classification

Unsupervised Classification dapat digunakan ketika tidak ada informasi tentang label data yang tersedia. Algoritma K-means clustering membagi data ke sejumlah *k* klaster sesuai dengan fitur yang ada. Pada kasus ini, wawancara dapat dibagi menjadi dua klaster, yaitu wawancara jujur dan wawancara palsu. Kelemahan metode *unsupervised classification* terhadap *supervised classification* secara umum adalah akurasinya.

Menggunakan software WEKA (Mark H dkk, 2009), pada tab clustering menggunakan simple *k* means clustering dimana *k*=2 diperoleh data bahwa *instance* yang salah klaster sebanyak 29,5%. Ini berarti bahwa akurasi yang didapatkan mencapai 70,5%. Gambar 7 memperlihatkan *output program*.

```
=== Model and evaluation on training set ===
Clustered Instances
0 379 ( 63%)
1 221 ( 37%)
Class attribute: interview
Classes to Clusters:
  0   1  <-- assigned to cluster
128 172 | REAL
251  49 | FALSIFIED
Cluster 0 <-- REAL
Cluster 1 <-- FALSIFIED
Incorrectly clustered instances :177.0 29.5 %
```

Gambar 7. Output 2-Means Clustering

4. Supervised Classification

Supervised classification dapat digunakan ketika ada informasi tentang pelabelan data untuk tiap *instance*. Dibandingkan *unsupervised classification*, *supervised classification* menawarkan tingkat akurasi yang lebih baik.

Metode umum untuk mengevaluasi *classifier* adalah *k-fold cross validation*, dimana data dibagi ke dalam sejumlah *k* bagian. Kemudian sejumlah *k*-1 bagian digunakan untuk melatih *classifier* dan 1 bagian untuk melakukan pengetesan.

Prosesnya kemudian dilakukan berulang sebanyak jumlah *k* dan akurasi final ditentukan dengan menghitung akurasi rata-rata dari setiap iterasi.

Classifier yang digunakan pada studi ini adalah regresi logistik. Regresi logistik sangat cocok sebagai *classifier binary*, sederhana, cepat, dan memiliki skalabilitas yang baik. Gambar 8 menunjukkan output software WEKA (Mark H dkk, 2009) menggunakan *10-fold cross validation*.

```
=== Summary ===
Correctly Classified Instances         88.5    %
Incorrectly Classified Instances       11.5    %
Kappa statistic                          0.77
Mean absolute error                      0.166
Root mean squared error                  0.2912
Relative absolute error                 33.1922 %
Root relative squared error             58.232  %
Total Number of Instances              600

=== Confusion Matrix ===
   a   b   <-- classified as
 272  28 |   a = REAL
  41 259 |   b = FALSIFIED
```

Gambar 8. Output Logistic Regression

IV. KESIMPULAN DAN SARAN

Kesimpulan

Berikut adalah kesimpulan dari hasil penelitian ini:

1. Fitur-fitur yang relevan dengan pemalsuan data didapatkan dari studi literatur dan dibagi menjadi dua kategori yaitu kategori data dan metadata. Untuk kategori metadata antara lain lokasi, *date stamp*, *time stamp*, durasi, dan *behavioral data* sedangkan untuk kategori data antara lain kesesuaian dengan hukum *benford*, variasi data, kombinasi jawaban yang jarang atau tidak lazim, dan unit yang tidak terisi.
2. Setelah dilakukan pemetaan dengan 15 sifat-sifat data palsu yang ada, ke-15 sifat data palsu tersebut dapat dicakup dengan fitur yang diusulkan.

3. Dari eksperimen yang dilakukan didapat kesimpulan tidak semua fitur efektif untuk mendeteksi pemalsuan data. fitur durasi dan tombol bantuan merupakan pemrediksi yang kuat dalam mendeteksi wawancara.
4. Klasifikasi otomatis dengan *supervised classification* memberikan akurasi yang lebih tinggi dari *unsupervised classification*. Namun *supervised classification* hanya dapat dilakukan dengan syarat terdapat informasi tentang data label.

Saran

Saran terkait penelitian ini antara lain:

1. Organisasi penyelenggara survei atau sensus sebaiknya merancang kuesioner dengan baik sehingga dapat lebih mengefektifkan waktu wawancara dan petugas tidak tertarik untuk melakukan pemalsuan data.
2. Sistem untuk monitoring dan pembuatan laporan dari sisi server dapat lebih dipelajari dan dikembangkan lagi.

DAFTAR PUSTAKA

A. Bennet, "Toward a solution of the “cheater problem” among parttime research investigators," Journal of Marketing 12 (4), pp. 470-474, 1948.

A. Kennickell, " Interviewers and data quality: Evidence from the 2001 survey of consumer finances," in Proceedings of the American Statistical Association (Survey Research Methods Section), 2002, p. 1807–1812.

A. Koch, "Fake Interview: Results of the Interviewer Control in ALLBUS 1994 ," ZUMA-Nachrichten 36, p. 89–105, 1995.

AAPOR. (2003) [Online].

http://www.aapor.org/Content/NavigationMenu/ResourcesforResearchers/falsification.pdf

Birnbaum B, Thessis: Algorithmic Approaches to Detecting Interviewer Fabrication in Surveys. University of Washington, 2012.

Birnbaum B., Gaetano B., Abraham D. F., Brian D., Anna R. K.. “Using Behavioral Data to Identify Interviewer Fabrication in Surveys”. Proceeding ACM CHI ‘13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Pages 2911-2920. 2013.

BPS, "Laporan Monitoring Kualitas Sensus Penduduk 2010," 2010.

BPS. (2011) Sirusa. [Online].

http://sirusa.bps.go.id/index.php?r=sd/view&kd=2164&th=2011

C. Hood and M. Bushery, "Getting more bang from the reinterviewer buck: Identifying ‘at risk’ interviewers," in In Proceedings of the American Statistical Association (Survey Research Methods Section), 1997., p. 820–824.

C. Lawrence and E. Love, "Characteristics of Falsified Interviews," Section of Survey Research Methods – JSM. , 2010.

C. Turner, J. Gribbe, A. Al-Tayyip, and J. Chromy, "Falsification in epidemiologic surveys: Detection and remediation (prepublication draft).," Technical Papers on Health and Behavior Measurement, No. 53., 2002.

I. Schreiner, K. Pennie, and J. Newbrough., "Interviewer falsification in census bureau surveys," in Proceedings of the American Statistical Association (Survey Research Methods Section), 1988, p. 491–496.

J. Murphy, R. Baxter, J. Eyerman, D. Cunningham, and J. Kennet, "A system for detecting interviewer falsification," the American Association for Public Opinion Research 59th Annual Conference, 2004.

J. Schrapler and G. Wagner, "Identification, characteristics and impact of faked interviews in surveys - an analysis by means of genuine fakes in the raw data of SOEP," IZA Discussion Paper Series, 969, 2003.

L. Crespi, "The cheater problem in polling," Public Opinion Quarterly 9 (4), p. 431–445, 1945.

M. Rita Thissen., "Computer Audio-Recorded Interviewing (CARI): A Tool for Monitoring Field Interviewers and Improving Field Data Collection," in Proceedings of Statistics Canada Symposium: Data Collection: Challenges, Achievements and New Directions, 2008.

Mark H., Eibe F., Geoffrey H., Bernhard P., Peter R., Ian H. W. "The WEKA Data Mining Software: An Update". SIGKDD Explorations, Volume 11, Issue 1. 2009

P. Harrisson, "A british view on "cheating"," Public Opinion Quarterly 11 (1), p. 172–173, 1947.

P. Kiecker and J. E. Nelson., "Do interviewers follow telephone survey instructions? ," Journal of the Market Research Society, p. 38:161–176, 1996.

R. Schnell, " The impact of counterfeit interviews on survey results," Journal of Sociology 20 (1), p. 25–35, 1991.

S. Bredl, N. Storfinger, and N. Menold, " A literature review of methods to detect fabricated survey data," ZEU, Universität Gießen, p. DiscussionPaer56, 2011.

S. Bredl, P. Winker, and K. Kotschau, "Technical Report 39: A statistical approach to detect cheating interviewers," 2008.

University of Maryland: Computer Assisted Surveys. [Online].

http://www.jpsm.umd.edu/surv400/notes/lecture%206%20Computer%20assisted%20surveys.pdf