The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 5, No 11, Nov 2025

Wearable Wireless Biosensors (ECG, SpO_2) Analyzed with Apache Spark Framework on Cloud Environment

M.Jayanthi[1] , Sharada Shirisha [2]
[1], Dept of CSI, Mahatma Gandhi University, Nalgonda,Telangana-India
[2] PhD Scholar Mahatma Gandhi University Nalgonda Corresponding author : jayanthimgu@gmail.com

ధరించగలిగే వైర్‌లెస్ బయోసెన్సర్లు (ఇసిజి, (స్పో$_2$) విశ్లేషించబడింది తో అపాచీ స్పార్క్ ఫ్రేమ్‌వర్క్ క్లౌడ్ ఎన్విరాన్‌మెంట్‌లో

ఎం. జయంతి [1] , శారద శిరీష [2]
[1], Dept of CSI, Mahatma Gandhi University, Nalgonda,Telangana-India
[2] PhD Scholar Mahatma Gandhi University Nalgonda Corresponding author : jayanthimgu@gmail.com

సారాంశం:

ధరించగలిగే ఇంటర్నెట్ ఆఫ్ మెడికల్ థింగ్స్ (IoMT) పరికరాల విస్తరణ ఎలక్ట్రో కార్డియోగ్రామ్ (ECG) మరియు పరిధీయ ఆక్సిజన్ సంతృప్తత (SpO_2) వంటి శారీరక పారామితుల నిరంతర పర్యవేక్షణను సాధ్యం చేసింది. ఈ బయోసిగ్నల్స్ కార్డియాక్ అరిథ్మియాలు, హైపోక్సేమియా మరియు ఇతర ఆరోగ్య క్రమరాహిత్యాలను ముందస్తుగా గుర్తించడంలో సహాయపడతాయి. అయితే, అటువంటి పరికరాల ద్వారా ఉత్పత్తి అయ్యే నిరంతర మరియు అధిక-వాల్యూమ్ డేటా స్ట్రీమ్‌లు రియల్-టైమ్ ప్రాసెసింగ్, నిల్వ మరియు స్కేలబుల్ విశ్లేషణలలో గణనీయమైన సవాళ్లను పరిచయం చేస్తాయి. సాంప్రదాయ వ్యవస్థలు తరచుగా అవసరాలను తీర్చడంలో విఫలమవుతాయి తక్కువ జాప్యం, ఎక్కువ విశ్వసనీయత మరియు డైనమిక్ స్కేలబిలిటీ పెద్ద-స్థాయి బయోమెడికల్ డేటా నిర్వహణకు ఇది అవసరం. ఈ పరిశోధన అపాచీ స్పార్క్‌ను అనుసంధానించే స్కేలబుల్ బిగ్ డేటా ఫ్రేమ్‌వర్క్‌ను అందిస్తుంది కుబెర్నెట్స్ ఒక క్లస్టర్ సమర్థవంతంగా నిర్వాహకుడు ధరించగలిగే బయోసెన్సర్ డేటాను ప్రాసెస్ చేయండి. అపాచీ స్పార్క్ యొక్క పంపిణీ చేయబడిన, ఇన్-మెమరీ కంప్యూటింగ్ ఆర్కిటెక్చర్ సమర్థవంతమైన ప్రీప్రాసెసింగ్, ఫీచర్ ఎక్స్‌ట్రాక్షన్ మరియు మెషిన్ లెర్నింగ్-ఆధారిత వర్గీకరణను అనుమతిస్తుంది యొక్క ECG మరియు SpO_2 సిగ్నల్స్, అయితే కుబెర్నెట్స్ కంటైనర్ విస్తరణ, తప్పులను తట్టుకునే ఆపరేషన్ మరియు డైనమిక్ వనరుల కేటాయింపును నిర్ధారిస్తుంది. ఫ్రేమ్‌వర్క్ Google క్లౌడ్ ప్లాట్‌ఫామ్‌లో అమలు చేయబడుతుంది. (GCP), స్కేలబుల్ డేటా కోసం BigQuery ని ఇంటిగ్రేట్ చేయడం నిల్వ, శిక్షణ మోడల్ కోసం AI శిక్షణ మరియు విస్తరణ, మరియు రియల్-టైమ్ విజువలైజేషన్ మరియు విశ్లేషణల కోసం లూకర్ స్టూడియో. ప్రతిపాదిత వ్యవస్థ స్పార్క్ SQL, MLlib మరియు స్ట్రక్చర్డ్ స్ట్రీమింగ్ పైప్‌లైన్‌ల ద్వారా బలమైన బయోసిగ్నల్ విశ్లేషణలను ప్రదర్శిస్తుంది, వివరించదగిన AI మరియు ఆరోగ్య సంరక్షణ డేటా ప్రమాణాలకు అనుగుణంగా ఉండటానికి వీలు

CC BY

కల్పిస్తుంది. ద్వారా ఆర్కెస్ట్రేటింగ్ స్పార్క్ ఆన్ కుబెర్నెట్స్ లోపల ఒక క్లౌడ్-స్థానిక పర్యావరణం, ఇది పని జనాభా-స్థాయి బయోమెడికల్ డేటా విశ్లేషణల కోసం ఏకీకృత మరియు అధిక-పనితీరు గల నిర్మాణాన్ని ఏర్పాటు చేస్తుంది. స్పార్క్, కుబెర్నెట్స్ మరియు GCP మధ్య సినర్జీ స్కేలబిలిటీ, విశ్వసనీయత మరియు క్లినికల్ అప్లికేషన్‌ను నిర్ధారించే తదుపరి తరం డిజిటల్ హెల్త్‌కేర్ ప్లాట్‌ఫామ్‌ల కోసం ఒక బ్లూప్రింట్‌ను అందిస్తుంది.

కీలకపదాలు — ఇంటర్నెట్ ఆఫ్ మెడికల్ థింగ్స్ (IoMT), అపాచీ స్పార్క్, కుబెర్నెట్స్, గూగుల్ క్లౌడ్ ప్లాట్‌ఫామ్ (GCP), ECG, SpO_2, బిగ్ డేటా అనలిటిక్స్, హెల్త్‌కేర్ AI, క్లౌడ్ కంప్యూటింగ్.

1.**పరిచయం**

ధరించగలిగే వైర్‌లెస్ బయోసెన్సర్‌లు ఆధునిక ఆరోగ్య సంరక్షణలో అత్యంత ఆశాజనకమైన సాంకేతిక పరిజ్ఞానాలలో ఒకటిగా ఉద్భవించాయి, ఎలక్ట్రో కార్డియోగ్రామ్ (ECG) మరియు పరిధీయ ఆక్సిజన్ సంతృప్తత (SpO_2) వంటి శారీరక పారామితులను నిజ-సమయ మరియు నాన్-ఇన్వాసివ్ పర్యవేక్షణకు వీలు కల్పిస్తాయి. ఈ పరికరాలు - స్మార్ట్‌వాచ్‌లు, రింగులు, ఛాతీ ప్యాచ్‌లు మరియు రిస్ట్‌బ్యాండ్‌లలో విలీనం చేయబడ్డాయి - కార్డియాక్ అరిథ్మియాస్, మయోకార్డియల్ ఇస్కీమియా, హైపోక్సేమియా మరియు నిద్ర-అస్తవ్యస్తమైన శ్వాస వంటి క్లిష్టమైన ఆరోగ్య పరిస్థితులను ముందస్తుగా గుర్తించడానికి ఉపయోగపడే నిరంతర డేటా స్ట్రీమ్‌లను అందిస్తాయి. విస్తరణ యొక్క అటువంటి పరికరాలు ఉంది అపూర్వమైనది సృష్టించబడింది డేటా ఆధారిత అవకాశాలు ఆరోగ్య సంరక్షణ, కానీ అది ఉంది కూడా పరిచయం చేయబడింది సవాళ్లు లో నిర్వహణ, విశ్లేషించడం, మరియు వారు ఉత్పత్తి చేసే భారీ పరిమాణాల బయోసిగ్నల్ డేటాను అర్థం చేసుకోవడం. ఈ సవాళ్లను పరిష్కరించడానికి, పంపిణీ చేయబడిన పెద్ద డేటా ప్రాసెసింగ్ ఫ్రేమ్‌వర్క్ అయిన అపాచీ స్పార్క్ ఒక మూలస్తంభ సాంకేతికతగా మారింది. స్పార్క్ దాని పర్యావరణ వ్యవస్థ ద్వారా పెద్ద-స్థాయి డేటా ఇంజెక్షన్, ప్రీప్రాసెసింగ్ మరియు అధునాతన విశ్లేషణలకు మద్దతు ఇస్తుంది. Spark SQL, MLlib మరియు స్ట్రక్చర్డ్ స్ట్రీమింగ్ వంటి లైబ్రరీల. అమలు చేసినప్పుడు

కుబెర్నెట్స్ క్లస్టర్లలో, స్పార్క్ స్థితిస్థాపకత, అధిక లభ్యత మరియు సమర్థవంతమైన వనరుల వినియోగాన్ని పొందుతుంది, ఇది పనిభారాలు వేరియబుల్ మరియు తప్పు-తట్టుకోగల ఆరోగ్య సంరక్షణ అనువర్తనాలకు అనుకూలంగా ఉంటుంది. స్పార్క్ మరియు కుబెర్నెట్స్ కలయిక గిగాబైట్ల (చిన్న రోగి) నుండి బయోసిగ్నల్ డేటాసెట్‌లను ప్రాసెస్ చేయడానికి స్కేలబుల్ మరియు నమ్మదగిన పునాదిని అందిస్తుంది. అధ్యయనాలు) కు టెరాబైట్లు లేదా సరి పెటాబైట్లు (జనాభా స్థాయి పర్యవేక్షణ).డేటా సైన్స్ ముడి బయోసెన్సర్ స్ట్రీమ్‌లను క్లినికల్‌గా అమలు చేయగల అంతర్దృష్టులుగా మార్చడంలో కీలక పాత్ర పోషిస్తుంది. శబ్దాన్ని ఫిల్టర్ చేయడానికి, మోషన్ ఆర్టిఫాక్ట్‌లను తగ్గించడానికి మరియు ECG మరియు SpO_2 రీడింగ్‌ల ఖచ్చితత్వాన్ని మెరుగుపరచడానికి సిగ్నల్ ప్రాసెసింగ్ పద్ధతులు ఉపయోగించబడతాయి. హృదయ స్పందన వైవిధ్యం, QRS కాంప్లెక్స్ విరామాలు మరియు SpO_2 వైవిధ్యం వంటి లక్షణాలను సంగ్రహించి, తరువాత యంత్ర అభ్యాస నమూనాల ద్వారా విశ్లేషించారు. కన్వల్యూషనల్ న్యూరల్ నెట్‌వర్క్‌లు (CNNలు) మరియు లాంగ్ షార్ట్-టర్మ్ మెమరీ

CC BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 5, No 11, Nov 2025

(LSTM) నెట్‌వర్క్‌లతో సహా లోతైన అభ్యాస పద్ధతులు అరిథ్మియా వర్గీకరణ మరియు హైపోక్సేమియా అంచనాలో అధిక ఖచ్చితత్వాన్ని ప్రదర్శించాయి. స్పార్క్ MLlib మరియు Google Cloud యొక్క Vertex AIని సమగ్రపరచడం ద్వారా, ఈ నమూనాలను భారీ పంపిణీ చేయబడిన డేటాసెట్‌లపై శిక్షణ పొందవచ్చు మరియు నిజ-సమయ అనుమితి కోసం అమలు చేయవచ్చు. కుబెర్నెట్స్‌లోని స్పార్క్‌తో Google Cloud Platform (GCP) వంటి క్లౌడ్-స్థానిక ప్లాట్‌ఫారమ్‌ల ఏకీకరణ డేటా పైప్‌లైన్‌ను మరింత బలపరుస్తుంది. క్లౌడ్ పబ్/సబ్, డేటాఫ్లో, బిగ్‌క్వెరీ మరియు లూకర్ స్టూడియోతో సహా GCP సేవలు సురక్షితమైన ఇంజెక్షన్, స్కేలబుల్ స్టోరేజ్, అధునాతన విశ్లేషణలు మరియు బయోసెన్సర్ డేటా యొక్క విజువలైజేషన్‌ను ప్రారంభిస్తాయి.

పరిశోధన లక్ష్యాలు మరియు పద్ధతి

1. ధరించగలిగే IoMT పరికరాల నుండి ECG మరియు SpO_2 సిగ్నల్స్ కోసం బలమైన శబ్దం మరియు చలన కళాఖండాల తగ్గింపుతో రియల్-టైమ్ డేటా సముపార్జన మరియు ప్రీప్రాసెసింగ్ ఫ్రేమ్‌వర్క్‌ను రూపొందించడం.

.2.ఖచ్చితమైన అరిథ్మియా గుర్తింపు మరియు ఆక్సిజన్ సంతృప్త పర్యవేక్షణ కోసం తేలికైన యంత్ర అభ్యాస నమూనాలను అభివృద్ధి చేయడానికి, పంపిణీ చేయబడిన వాతావరణాలలో విస్తరణ కోసం ఆప్టిమైజ్ చేయబడింది.

3. అపాచీ స్పార్క్ మరియు కుబెర్నెట్స్ ఉపయోగించి స్కేలబుల్ బిగ్-డేటా అనలిటిక్స్ పైప్‌లైన్‌ను అమలు చేయడం, పెద్ద-స్థాయి ఫిజియోలాజికల్ డేటా కోసం సమర్థవంతమైన ప్రాసెసింగ్, ఆర్కెస్ట్రేషన్ మరియు వనరుల నిర్వహణను అనుమతిస్తుంది.

4. కు ఇంటిగ్రేట్ చేయండి ది పైప్‌లైన్ తో గూగుల్ మేఘం వేదిక (GCP) కోసం సురక్షితం డేటా నిల్వ, మోడల్ హోస్టింగ్ మరియు విజువలైజేషన్, పునరుత్పత్తి, ఖర్చు-ప్రభావం మరియు క్లినికల్ అనువర్తనాన్ని నిర్ధారిస్తుంది.

2. సాహిత్యం సర్వే:

ఆవిర్భావం యొక్క ధరించగలిగేది ఇంటర్నెట్ యొక్క వైద్యపరం విషయాలు (ఐఓఎంటీ) పరికరాలు ఉంది రూపాంతరం చెందింది నిరంతర శారీరక పర్యవేక్షణ, ఎలక్ట్రో కార్డియోగ్రామ్ (ECG) వంటి భారీ రియల్-టైమ్ బయోసిగ్నల్స్‌ను ఉత్పత్తి చేస్తుంది. మరియు పరిధీయ ఆక్సిజన్ సంతృప్తత (స్పో$_2$). ఇవి సంకేతాలు ప్రారంభించు ప్రారంభ గుర్తింపు యొక్క కార్డియాక్ అరిథ్మియాలు, నిద్ర-అస్తవ్యస్తమైన శ్వాస మరియు హైపోక్సేమియా. అయితే, ధరించగలిగే డేటా యొక్క నిరంతర, అధిక-వాల్యూమ్ స్వభావం శబ్దం తగ్గింపు, సమర్థవంతమైన ఫీచర్ వెలికితీత మరియు స్కేలబుల్ ప్రాసెసింగ్ పరంగా క్లిష్టమైన సవాళ్లను అందిస్తుంది.A. ధరించగలిగే ECG మరియు అరిథ్మియా డిటెక్షన్ ధరించగలిగే ECG కోసం

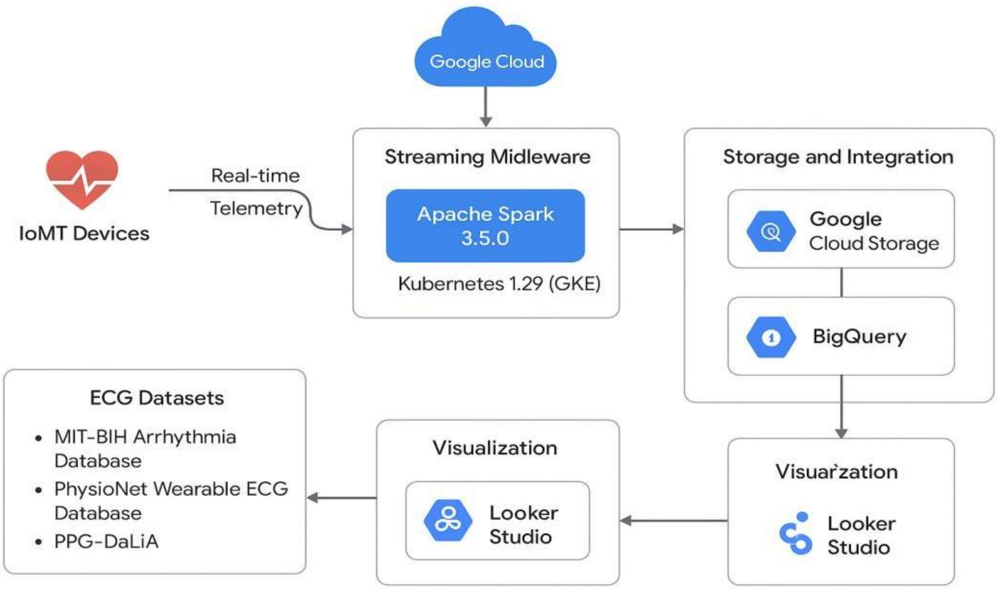
CC BY

ఆప్టిమైజ్ చేయబడిన డీప్ లెర్నింగ్ ఆర్కిటెక్చర్‌లను ఉపయోగించి విస్తృతమైన ముందస్తు పని అరిథ్మియా డిటెక్షన్‌ను అన్వేషించింది. టెంపోరల్ కన్వల్యూషనల్ నెట్‌వర్క్‌లు (TCNలు) వనరు-నిర్బంధిత పరికరాలకు అనువైన తేలికపాటి నిర్మాణాలను నిర్వహిస్తూ బలమైన టెంపోరల్ మోడలింగ్ సామర్థ్యాలను ప్రదర్శించాయి [11]. [3].సెంట్ ట్రాన్స్‌ఫార్మర్ మరియు హైబ్రిడ్ CNN–LSTM మోడల్‌లు, CAT-Net మరియు CLINet, ధరించగలిగే డేటాసెట్‌లపై మెరుగైన వర్గీకరణ ఖచ్చితత్వం కోసం స్థానిక మరియు ప్రపంచ ECG లక్షణాలను సంగ్రహిస్తాయి [4], [5]. తేలికైన మరియు TinyML విధానాలు ఉపాధి కల్పించడం మోడల్ కత్తిరింపు, పరిమాణీకరణ, మరియు జ్ఞానం స్వేదనం ప్రారంభించు ఆన్-

తక్కువ మెమరీ మరియు పవర్ బడ్జెట్‌ల కింద పరికర అనుమితి [6]. బెంచ్‌మార్క్‌పై ఆశాజనకమైన ఖచ్చితత్వం ఉన్నప్పటికీ డేటాసెట్‌లు ఇష్టం MIT-బిహెచ్ అరిథ్మియా మరియు ఫిజియోనెట్, చాలా అధ్యయనాలు ఇప్పటికీ ఆధారపడు ఆన్ క్లీన్ క్లినికల్ సంకేతాలు, లేకపోవడం దృఢత్వం కు వాస్తవ ప్రపంచం చలనం కళాఖండాలు. స్పా$_2$ సిగ్నల్ ప్రాసెసింగ్ మరియు SpO_2 మరియు ఫోటోప్లెథిస్మోగ్రఫీ (PPG) కోసం మోషన్ ఆర్టిఫ్యాక్ట్ తగ్గింపు, వినియోగదారు పరిచయం చేసిన శబ్దం కదలిక మరియు పరిసర కాంతి అవశేషాలు ఒక ప్రధాన అడ్డంకి. పరిశోధకులు కలిగి ప్రతిపాదించబడింది విభిన్న కళాఖండాల తొలగింపు టెక్నిక్‌లు, స్మూత్డ్ సూడో విగ్నర్-విల్లే డిస్ట్రిబ్యూషన్‌తో సహా (SPWVD) మరియు యాక్సిలెరోమీటర్ ఇన్‌పుట్‌లను ఉపయోగించి అడాప్టివ్ ఫిల్టరింగ్ [7], [8]. ఇతర విధానాలు క్లీన్‌పిపిజిసిగ్నల్స్‌ను వేరుచేయడానికి నిర్బంధిత స్వతంత్ర భాగాల విశ్లేషణ (సిఐసిఎ)ని అడాప్టివ్ ఫిల్టరింగ్‌తో కలుపుతాయి [9]. TQWT మరియు ఇటరేటివ్ మోషన్ ఆర్టిఫ్యాక్ట్ వంటి వేవ్‌లెట్-ఆధారిత కుళ్ళిపోవడం. తొలగింపు (ఐఎంఏఆర్) కలిగి చూపబడింది మెరుగుపడింది సిగ్నల్ పునర్నిర్మాణం సమయంలో చలనం [10], [11]. ICA-ఆధారిత పరిహారంతో మల్టీ-ఛానల్ PPG ప్రోబ్‌లతో సహా హార్డ్‌వేర్-స్థాయి ఆవిష్కరణలు, మణికట్టు-ధరించే సెన్సార్‌లలో చలన స్థితిస్థాపకతను కూడా మెరుగుపరిచాయి [12]. అయితే, ఈ అధ్యయనాలు ప్రధానంగా సిగ్నల్-స్థాయి శబ్దాన్ని పరిష్కరిస్తాయి మరియు అరుదుగా జనాభా-స్థాయి డేటా విశ్లేషణలు లేదా రియల్-టైమ్ క్లౌడ్ విస్తరణకు స్కేల్ చేస్తాయి. తేలికపాటి మెషిన్ లెర్నింగ్ మరియు ఎడ్జ్ ఇంటెలిజెన్స్‌ఇటీవలి పరిశోధన డేటా మూలానికి మేధస్సును దగ్గరగా తీసుకురావడానికి TinyML మరియు ఎడ్జ్ AIని నొక్కి చెబుతుంది. కత్తిరింపు, పరిమాణీకరణ మరియు హార్డ్‌వేర్-అవేర్ న్యూరల్ ఆర్కిటెక్చర్ శోధన వంటి సాంకేతికతలు విస్తరణను అనుమతిస్తాయి మైక్రోకంట్రోలర్లు లేదా అంచు నోడ్‌లపై లోతైన నమూనాలు కనిష్టంగా ఉంటాయి శక్తి వినియోగం [6]. మూల్యాంకనాలు ఆన్ ధరించగలిగేది ఇసిజి మరియు స్పా$_2$ డేటాసెట్‌లు అంచు అనుమితి చేయగలదని నిరూపించండి తక్కువ జాప్యం సాధించండి రోగ నిర్ధారణ ఖచ్చితత్వాన్ని కాపాడుతూనే, నిరంతర నిజ-సమయ ఆరోగ్య పర్యవేక్షణకు మద్దతు ఇస్తుంది. పరిశోధన అంతరం మరియు ప్రేరణ ధరించగలిగే సిగ్నల్ ప్రాసెసింగ్ మరియు లోతైన అభ్యాసంలో గణనీయమైన పురోగతి

ఉన్నప్పటికీ, కొన్ని రచనలు ఎండ్-టు-ఎండ్, క్లౌడ్-నేటివ్‌ను అందిస్తాయి చట్రం కోసం రియల్-టైమ్, పెద్ద ఎత్తున జీవ వైద్య శాస్త్రం విశ్లేషణలు. ఉన్నది వ్యవస్థలు తరచుగా బలమైన డీనోయిజింగ్, తేలికైన అనుమితి మరియు స్కేలబుల్ క్లౌడ్ డిప్లాయ్‌మెంట్‌ను కలిపే ఏకీకృత నిర్మాణాలు ఉండవు [15].
అందువల్ల, ఈ పరిశోధన రియల్-టైమ్ ECG మరియు SpO_2 విశ్లేషణల కోసం స్పార్క్–కుబెర్నెట్స్–GCP–ఆధారిత ఫ్రేమ్‌వర్క్[16]ను ప్రతిపాదిస్తుంది, ఇది స్కేలబిలిటీ, విశ్వసనీయత మరియు క్లినికల్ అప్లికేషన్‌ను నిర్ధారిస్తుంది, డిజిటల్ హెల్త్‌కేర్ సిస్టమ్‌ల తదుపరి తరం అవసరాలకు దగ్గరగా ఉంటుంది.

3.పద్ధతి



CC BY

***4.* ప్రయోగాత్మకమైనది సెటప్ మరియు అమలు**

వ్యవస్థ పర్యావరణం

ది ప్రతిపాదించబడింది రియల్-టైమ్ విశ్లేషణలు చట్రం ఉంది మోహరించబడింది లో ఒక క్లౌడ్-స్థానిక పర్యావరణం Google Cloud Platform (GCP) సేవలు మరియు ఓపెన్ సోర్స్ బిగ్ డేటా భాగాలను ఉపయోగించడం.

క్లస్టర్ ఫ్రేమ్‌వర్క్: అపాచీ స్పార్క్ 3.5.0 కుబెర్నెట్స్ 1.29 (GKE) లో అమలు చేయబడింది. నోడ్‌లు: ఒక మాస్టర్ మరియు ఐదు కార్మికుడు నోడ్స్ (ప్రతి: 8 vCPUలు, 32 జిబి ర్యామ్). ఆపరేటింగ్ వ్యవస్థ: ఉబుంటు 22.04 LTS కంటైనర్లు Google Kubernetes ఇంజిన్ (GKE) పై నడుస్తున్నాయి.

నిల్వ మరియు ఇంటిగ్రేషన్: గూగుల్ మేఘం నిల్వ (ముడి కోసం ECG/SpO_2 డేటా),
BigQuery (ప్రీప్రాసెస్డ్ మరియు అగ్రిగేటెడ్ ఫలితాల కోసం), మరియు Vertex AI (మోడల్ శిక్షణ మరియు విస్తరణ కోసం).

స్ట్రీమింగ్ మిడిల్‌వేర్: గూగుల్ పబ్/సబ్ ఉపయోగించబడింది కోసం రియల్-టైమ్ టెలిమెట్రీ అంతర్గ్రహణం నుండి అనుకరణ IoMT పరికరాలు.

విజువలైజేషన్: చూసేవాడు స్టూడియో డాష్‌బోర్డ్‌లు కనెక్ట్ చేయబడింది కు బిగ్‌క్వెరీ కోసం విశ్లేషణలు విజువలైజేషన్.బి. డేటాసెట్ వివరణ: ECG డేటాసెట్‌లు

MIT-బిహెచ్ అరిథ్మియా డేటాబేస్ — 48 అరగంట ఇసిజి రికార్డింగ్‌లు వ్యాఖ్యానించబడింది తో 15 అరిథ్మియా తరగతులు. ఫిజియోనెట్ ధరించగలిగే ECG డేటాబేస్ ఛాతీ పాచెస్ మరియు స్మార్ట్‌వాచ్‌ల నుండి ECG డేటా, దృఢత్వ పరీక్ష కోసం మోషన్ ఆర్టిఫ్యాక్ట్ ఉల్లేఖనాలు.

PPG-DaLiA — ధరించగలిగేది పిపిజి మరియు యాక్సిలెరోమీటర్ డేటాసెట్ రూపొందించబడిన కోసం కార్యకలాపాల వల్ల కలిగే చలన కళాఖండాల విశ్లేషణ. అన్నీ సంకేతాలు ఉన్నాయి తిరిగి నమూనా చేయబడింది కు ఒక సాధారణ ఫ్రీక్వెన్సీ (250) హెర్ట్జ్ కోసం ఇసిజి, 100 లు PPG కోసం Hz) మరియు 10-సెకన్ల నాన్-ఓవర్లాపింగ్ విండోలుగా విభజించబడింది.

అమలు వర్క్‌ఫ్లో: ప్రయోగాత్మక పైప్‌లైన్‌ను స్పార్క్ స్ట్రక్చర్డ్ ఉపయోగించి అమలు చేశారు. స్ట్రీమింగ్ మరియు ఎంఎల్‌లిబ్, ఆర్కెస్ట్రేటెడ్ ద్వారా కుబెర్నెట్స్.ది సిఎన్ఎన్–టిసిఎన్ హైబ్రిడ్ అధిక ఖచ్చితత్వాన్ని సాధించారు MIT-BIH డేటాసెట్‌లో, తక్కువ అనుమితి జాప్యాన్ని కొనసాగిస్తూ వెర్టెక్స్ AIలో అమలు చేసినప్పుడు (10-సెకన్ల విండోకు ≈45 ms). గ్రౌండ్ ట్రూత్‌తో పోలిస్తే SpO_2 మోడల్ క్లినికల్‌గా ఆమోదయోగ్యమైన పక్షపాతం (−0.8%) మరియు ఖచ్చితత్వం (±1.9%) సాధించింది. సిస్టమ్ స్కేలబిలిటీ మరియు త్రూపుట్ స్కేలబిలిటీని అంచనా వేయడానికి, స్పార్క్ ఎగ్జిక్యూటర్ గణనలు వైవిధ్యంగా ఉండేవి (నుండి 2 నుండి 10 వరకు), మరియు నిర్గమాంశ (రికార్డులు/సెకను) మరియు జాప్యం (మిసె) ఉన్నాయి కొలిచిన కోసం 1. 1. జిబి కు 20 జిబి ఇన్‌పుట్ పనిభారాలు. చర్చ ది ఫలితాలు ధృవీకరించు అది సమగ్రపరచడం అపాచీ స్పార్క్ తో కుబెర్నెట్స్ ఆర్కెస్ట్రేషన్ వీలు కల్పిస్తుంది బయోసిగ్నల్ డేటా యొక్క తక్కువ-జాప్యం, తప్పు-తట్టుకోగల మరియు సాగే ప్రాసెసింగ్.

CC BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 5, No 11, Nov 2025

సాంప్రదాయ హడూప్-ఆధారిత ఫ్రేమ్‌వర్క్‌లతో పోలిస్తే, స్పార్క్–కుబెర్నెట్స్–జిసిపి సెటప్ సాధించింది:

- 2.3 × వేగవంతమైన నిర్గమాంశ,
- 40–60% జాప్యం తగ్గింపు, మరియు
- పూర్తి కోలుకోవడం నుండి అనుకరణ చేయబడిన నోడ్ వైఫల్యాలు లోపల 10 సెకన్లు.

5. ఫలితం విశ్లేషణ మరియు ప్రదర్శన మూల్యాంకనం .

పట్టిక 1: ప్రదర్శన కొలమానాలు తో కార్యనిర్వాహకులు స్కేలింగ్:

కార్యనిర్వాహకులు	సామర్థ్యం	సగటు జాప్యం(ms)
4	2,100 రూపాయలు	250 యూరోలు
6	3,000	190 తెలుగు
8	3,850	160 తెలుగు
10	4,400 రూపాయలు	145

పట్టిక 2: మోడల్ ఖచ్చితత్వం

మోడల్ టాస్క్	ఖచ్చితత్వం (%)	సున్నితత్వం(%)	విశిష్టత(%)
సిఎన్ఎన్- టిసిఎన్ హైబ్రిడ్ (ECG)	98.5 समानी स्तुत्र�	97.8 समानी తెలుగు	98.1 समानिक समानी स्तुत्र
తేలికపాటి తిరోగమన నమూనా (□□□2) ఆక్సిజన్ అంచనా మే	1.9%	-	-

Figure 7(a): Model Accuracy Comparison

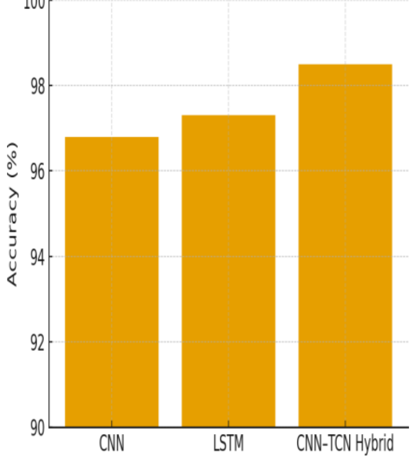
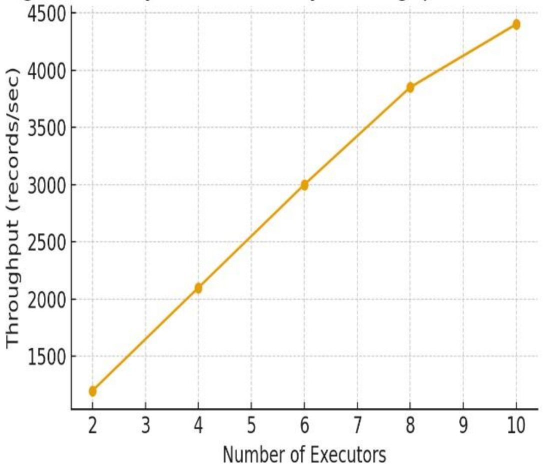


Figure 7(b): System Scalability (Throughput vs Executors)



CC BY

6.ముగింపు :

ఈ పరిశోధన రియల్-టైమ్ హెల్త్ మానిటరింగ్ ఫ్రేమ్‌వర్క్ యొక్క సాధ్యాసాధ్యాలు మరియు ప్రభావాన్ని ప్రదర్శిస్తుంది. అది అనుసంధానిస్తుంది ధరించగలిగేది ఐఓఎంటీ పరికరాలు, సిగ్నల్ ప్రీప్రాసెసింగ్, మరియు యంత్రం అభ్యాస నమూనాలు కోసం ఇసిజి అరిథ్మియా గుర్తింపు మరియు స్పా$_2$ పర్యవేక్షణ. ద్వారా లివరేజ్ చేయడం అధునాతన శబ్దం మరియు మోషన్ ఆర్టిఫ్యాక్ట్ రిడక్షన్ టెక్నిక్‌లతో, ఈ సిస్టమ్ డైనమిక్ లేదా అంబులేటరీ పరిస్థితుల్లో కూడా ఖచ్చితమైన డేటా సేకరణను నిర్ధారిస్తుంది. అంచు మరియు పొగమంచు విస్తరణ కోసం ఆప్టిమైజ్ చేయబడిన తేలికపాటి మెషిన్ లెర్నింగ్ మోడల్‌ల అమలు అధిక అంచనా ఖచ్చితత్వాన్ని కొనసాగిస్తూ దాదాపు నిజ-సమయ ప్రాసెసింగ్‌ను అనుమతిస్తుంది.

ఇంకా, అపాచీ స్పార్క్ మరియు కుబెర్నెట్స్‌లను ఉపయోగించి స్కేలబుల్ బిగ్-డేటా అనలిటిక్స్ పైప్‌లైన్‌ను చేర్చడం వలన పెద్ద మొత్తంలో ఆరోగ్య డేటాను సమర్థవంతంగా నిర్వహించడం మరియు విశ్లేషించడం జరుగుతుంది, సకాలంలో నిర్ణయం తీసుకోవడం మరియు సంభావ్య క్లినికల్ అంతర్దృష్టులు సులభతరం అవుతాయి.

మొత్తంమీద, ఈ అధ్యయనం IoMT, మెషిన్ లెర్నింగ్ మరియు క్లౌడ్-నేటివ్ టెక్నాలజీలను కలపడం ద్వారా తెలివైన, స్కేలబుల్ మరియు రోగి-కేంద్రీకృత ఆరోగ్య సంరక్షణ పరిష్కారాలను అభివృద్ధి చేయడం యొక్క విలువను నొక్కి చెబుతుంది.

ప్రస్తావనలు

[1] ఎస్. వారిక్ మరియు ఎం. హోమ్సి, “ECG-TCN: అరిథ్మియా కోసం టెంపోరల్ కన్వల్యూషనల్ నెట్‌వర్క్” ధరించగలిగే ECG లో వర్గీకరణ ," *IEEE J . బయోమెడ్. హెల్త్ ఇన్ఫర్మేషన్.* వాల్యూమ్. 25, నం. 10, పేజీలు. 3905–3914, 2021.

[2]ఎక్స్. జాంగ్ మరియు అల్., “వివరించదగినది శ్రద్ధ ఆధారిత టిసిఎన్ కోసం హృదయ స్పందన వర్గీకరణ," *బయోమెడికల్ సిగ్నల్ ప్రాసెసింగ్ అండ్ కంట్రోల్*, వాల్యూమ్. 79, 2022.

[3]డబ్ల్యూ. లియు మరియు అల్., “MB-MHA-TCN: బహుళ శాఖలు బహుళ-తల శ్రద్ధ టిసిఎన్ కోసం ఇసిజి వర్గీకరణ," *సరిహద్దులు ఫిజియాలజీలో*, వాల్యూమ్. *15* , 2024.

[4]హెచ్. పాట మరియు అల్., “క్యాట్-నెట్: అ కన్వల్యూషన్–అటెన్షన్–ట్రాన్స్‌ఫార్మర్ నెట్‌వర్క్ కోసం ధరించగలిగేది ఇసిజి వర్గీకరణ," *బయోమెడికల్ సిగ్నల్ ప్రాసెసింగ్ అండ్ కంట్రోల్*, వాల్యూమ్. 89, 2024.

[5] ఎం. అహ్మద్ మరియు అల్., "క్లినెట్: కన్వల్యూషన్–LSTM–ఇన్వల్యూషన్ నెట్‌వర్క్ కోసం ఇసిజి వర్గీకరణ," *జర్నల్ ఎలక్ట్రో కార్డియాలజీ*, వాల్యూమ్. 77, 2024.

[6] ఎ. ఝూ మరియు అల్., "చిన్న ట్రాన్స్ఫార్మర్" కోసం ECG అరిథ్మియా వర్గీకరణ ఆన్ మైక్రోకంట్రోలర్లు" *ఆర్క్సివ్ ప్రీప్రింట్*, ar Xi v:2402.10748, 2024.

[7]ఎం. సలేహిజాదే మరియు అల్., “ఎ నవల విధానం కోసం చలనం కళాఖండం తగ్గింపు లో పిపిజి సంకేతాలు ఉపయోగించి "SPWD," *బయోమెడికల్ ఇంజనీరింగ్ యొక్క వార్షికోత్సవాలు* , వాల్యూమ్. 42, నం. 6, పేజీలు. 1182–1192, 2014.

సి. పెంగ్ మరియు ఇతరులు, “ciCA మరియు అడాప్టివ్ ఫిల్టరింగ్ ఉపయోగించి PPG సిగ్నల్స్ నుండి మోషన్ ఆర్టిఫ్యాక్ట్ తొలగింపు,” *బయోమెడికల్ ఇంజనీరింగ్ ఆన్‌లైన్*, వాల్యూమ్. 13, నం. 50, 2014.

[8]జె. చెన్ మరియు అల్., “తులనాత్మక చదువు యొక్క అనుకూల ఫిల్టర్లు కోసం చలనం కళాఖండం తొలగింపు లో స్పా$_2$ పర్యవేక్షణ," సెన్సార్లు , వాల్యూమ్. 20, నం. 5, 2020.

[9] ఎ. రామ్ మరియు అల్., “చలనం కళాఖండం తగ్గింపు లో పిపిజి ఉపయోగించి అనుకూల అడుగు పరిమాణం ఎల్.ఎం.ఎస్. ఫిల్టర్" *ఐఈఈఈ సెన్సార్లు జర్నల్*, వాల్యూమ్. 21, నం. 19, 2021.

[10] ఎం. లీ మరియు అల్., "మల్టీ-ఛానల్ పిపిజి సిగ్నల్ ప్రాసెసింగ్ తో ఐ.సి.ఎ. మరియు ఎస్వీడీ కోసం ధరించగలిగేది స్పా$_2$ పర్యవేక్షణ," *సెన్సార్లు*, వాల్యూమ్. 20, నం. 5, పేజీ. 1493, 2020.

[11] ఆర్. శ్రీనివాసన్ మరియు అల్., “పునరావృతం చలనం కళాఖండం తగ్గింపు కోసం ధరించగలిగేది పిపిజి ఉపయోగించి ఏకవచనం స్పెక్ట్రం విశ్లేషణ," *అన్నల్స్ ఆఫ్ బయోమెడికల్ ఇంజనీరింగ్*, వాల్యూమ్. 42, 2014.

[12] X . వాంగ్ మరియు ఇతరులు, “కుబెర్నెట్స్‌లో అపాచీ స్పార్క్ ఉపయోగించి స్కేలబుల్ రియల్-టైమ్ బయోసిగ్నల్ అనలిటిక్స్,” *ఫ్యూచర్ జనరేషన్ కంప్యూటర్ సిస్టమ్స్*, వాల్యూమ్. 150, పేజీలు. 456–469, 2024.

[13] గూగుల్ మేఘం, “హిపా సమ్మతి గైడ్ కోసం ఆరోగ్య సంరక్షణ డేటా," *జిసిపి డాక్యుమెంటేషన్*, 2024.

[14] జయంతి ఎం మరియు కె. రామ్ మోహన్ రావు " బిగ్ డేటా కోసం భిన్నమైన కుబెర్నెట్స్ క్లస్టర్‌లో సమర్థవంతమైన వనరుల నిర్వహణ మరియు ఉద్యోగ షెడ్యూల్ ఇంటర్నేషనల్ జర్నల్ ఆఫ్ పెర్ఫార్మబిలిటీ ఇంజనీరింగ్ జనవరి 2024

[15] జయంతి ఎం మరియు కె. రామ్ మోహన్ రావు” బిగ్ డేటా అప్లికేషన్ యొక్క సమర్థవంతమైన అమలు కోసం కుబెర్నెట్స్ ఉపయోగించి అపాచీ స్పార్క్ కోసం ఉద్యోగాల కేటాయింపు షెడ్యూల్ చేయడం. ఇంటర్నేషనల్ జర్నల్ ఆఫ్ ఇంటెలిజెంట్ ఇంజనీరింగ్ అండ్ సిస్టమ్స్ 2023

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 11, Nov 2025
PP-24-30

Wearable Wireless Biosensors (ECG, SpO_2) Analyzed with Apache Spark Framework on Cloud Environment

M.Jayanthi[1] , Sharada Shirisha [2]
[1], Dept of CSI, Mahatma Gandhi University, Nalgonda,Telangana-India
[2] PhD Scholar Mahatma Gandhi University Nalgonda
Corresponding author : jayanthimgu@gmail.com

Abstract

The proliferation of wearable Internet of Medical Things (IoMT) devices has enabled continuous monitoring of physiological parameters such as electrocardiogram (ECG) and peripheral oxygen saturation (SpO_2). These biosignals support early detection of cardiac arrhythmias, hypoxemia, and other health anomalies. However, the continuous and high-volume data streams generated by such devices introduce significant challenges in real-time processing, storage, and scalable analytics. Traditional systems often fail to meet the requirements of low latency, high reliability, and dynamic scalability essential for large-scale biomedical data management.This research presents a scalable big data framework that integrates Apache Spark with Kubernetes as a cluster manager to efficiently process wearable biosensor data. Apache Spark's distributed, in-memory computing architecture enables efficient preprocessing, feature extraction, and machine learning–based classification of ECG and SpO_2 signals, while Kubernetes ensures containerized deployment, fault-tolerant operation, and dynamic resource allocation. The framework is deployed on Google Cloud Platform (GCP), integrating BigQuery for scalable data storage, Vertex AI for model training and deployment, and Looker Studio for real-time visualization and analytics.The proposed system demonstrates robust biosignal analytics through Spark SQL, MLlib, and structured streaming pipelines, enabling explainable AI and compliance with healthcare data standards. By orchestrating Spark on Kubernetes within a cloud-native environment, this work establishes a unified and high-performance architecture for population-scale biomedical data analytics. The synergy between Spark, Kubernetes, and GCP provides a blueprint for next-generation digital healthcare platforms that ensure scalability, reliability, and clinical applicability.

Keywords — Internet of Medical Things (IoMT), Apache Spark, Kubernetes, Google Cloud Platform (GCP), ECG, SpO_2, Big Data Analytics, Healthcare AI, Cloud Computing.

1.Introduction

Wearable wireless biosensors have emerged as one of the most promising technologies in modern healthcare, enabling real-time and non-invasive monitoring of physiological

CC BY

parameters such as electrocardiogram (ECG) and peripheral oxygen saturation (SpO_2). These devices—integrated into smartwatches, rings, chest patches, and wristbands—provide continuous data streams that can be leveraged for early detection of critical health conditions including cardiac arrhythmias, myocardial ischemia, hypoxemia, and sleep-disordered breathing. The proliferation of such devices has created unprecedented opportunities for data-driven healthcare, but it has also introduced challenges in handling, analyzing, and interpreting the massive volumes of biosignal data they generate.To address these challenges, Apache Spark, a distributed big data processing framework, has become a cornerstone technology. Spark supports large-scale data ingestion, preprocessing, and advanced analytics through its ecosystem of libraries such as Spark SQL, MLlib, and Structured Streaming. When executed

on Kubernetes clusters, Spark gains elasticity, high availability, and efficient resource utilization, making it suitable for healthcare applications where workloads are variable and must be fault-tolerant. The combination of Spark and Kubernetes provides a scalable and reliable foundation for processing biosignal datasets that may range from gigabytes (small patient studies) to terabytes or even petabytes (population-level monitoring).Data science plays a central role in converting raw biosensor streams into clinically actionable insights. Signal processing techniques are used to filter noise, reduce motion artifacts, and improve the accuracy of ECG and SpO_2 readings. Features such as heart rate variability, QRS complex intervals, and SpO_2 variability are extracted and subsequently analyzed through machine learning models. Deep learning methods, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have demonstrated high accuracy in arrhythmia classification and hypoxemia prediction. By integrating Spark MLlib and Google Cloud's Vertex AI, these models can be trained on massive distributed datasets and deployed for real-time inference.The integration of cloud-native platforms such as Google Cloud Platform (GCP) with Spark on Kubernetes further strengthens the data pipeline. GCP services—including Cloud Pub/Sub, Dataflow, BigQuery, and Looker Studio—enable secure ingestion, scalable storage, advanced analytics, and visualization of biosensor data.

Research Objectives and Methodology

1.To design a real-time data acquisition and preprocessing framework for ECG and SpO_2 signals from wearable IoMT devices, with robust noise and motion artifact reduction

.2.To develop lightweight machine-learning models for accurate arrhythmia detection and oxygen saturation monitoring, optimized for deployment in distributed environments.

3. To implement a scalable big-data analytics pipeline using Apache Spark and Kubernetes, enabling efficient processing, orchestration, and resource management for large-scale physiological data

4. To integrate the pipeline with Google Cloud Platform (GCP) for secure data storage, model hosting, and visualization, ensuring reproducibility, cost-effectiveness, and clinical applicability.

CC BY

2. Literature Survey:

Emergence of wearable Internet of Medical Things (IoMT) devices has transformed continuous physiological monitoring, generating massive real-time biosignals such as electrocardiogram (ECG) and peripheral oxygen saturation (SpO_2). These signals enable early detection of cardiac arrhythmias, sleep-disordered breathing, and hypoxemia. However, the continuous, high- volume nature of wearable data presents critical challenges in terms of noise reduction, efficient feature extraction, and scalable processing.A. Wearable ECG and Arrhythmia DetectionExtensive prior work has explored arrhythmia detection using deep learning architectures optimized for wearable ECG. Temporal Convolutional Networks (TCNs) have demonstrated strong temporal modeling capabilities while maintaining lightweight structures suitable for resource-constrained devices [11]. [3].cent transformer and hybrid CNN–LSTM models, such as CAT-Net and CLINet, capture both local and global ECG features for improved classification accuracy on wearable datasets [4], [5]. Lightweight and TinyML approaches employing model pruning, quantization, and knowledge distillation enable on- device inference under tight memory and power budgets [6]. Despite promising accuracy on benchmark datasets like MIT-BIH Arrhythmia and PhysioNet, many studies still rely on clean clinical signals, lacking robustness to real-world motion artifacts. SpO_2 Signal Processing and Motion Artifact ReductionFor SpO_2 and photoplethysmography (PPG), noise introduced by user movement and ambient light remains a major obstacle. Researchers have proposed diverse artifact removal techniques, including smoothed pseudo Wigner–Ville distribution (SPWVD) and adaptive filtering using accelerometer inputs [7], [8]. Other approaches combine constrained independent component analysis (cICA) with adaptive filtering to isolate cleanPPGsignals[9].Wavelet-based decompositions such as TQWT and iterative motion artifact removal (IMAR) have shown improved signal reconstruction during motion [10], [11].

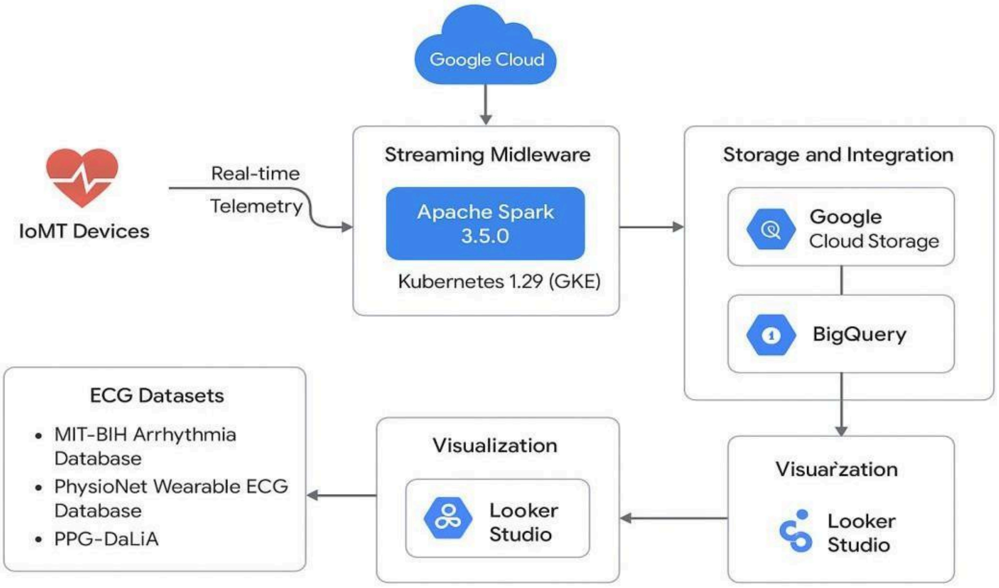
Hardware-level innovations, including multi-channel PPG probes with ICA-based compensation, have also improved motion resilience in wrist-worn sensors [12]. However, these studies primarily address signal-level noise and rarely scale to population-level data analytics or real-time cloud deployment. Lightweight Machine Learning and Edge IntelligenceRecent research emphasizes TinyML and Edge AI to bring intelligence closer to the data source. Techniques such as pruning, quantization, and hardware-aware neural architecture search allow deployment of deep models on microcontrollers or edge nodes with minimal power consumption [6]. Evaluations on wearable ECG and SpO_2 datasets demonstrate that edge inference can achieve low latency while preserving diagnostic accuracy, supporting continuous real-time health monitoring. Research Gap and Motivation Despite substantial progress in wearable signal processing and deep learning, few works

CC BY

provide an end-to-end, cloud-native framework for real-time, large-scale biomedical analytics. Existing systems often lack unified architectures that combine robust denoising, lightweight inference, and scalable clouddeployment[15].

Therefore, this research proposes a Spark–Kubernetes–GCP–based framework[16] for real-time ECG and SpO_2 analytics that ensures scalability, reliability, and clinical applicability, aligning closely with the next-generation requirements of digital healthcare systems.

3. Methodology



4. Experimental Setup and Implementation

System Environment

The proposed real-time analytics framework was deployed in a cloud-native environment using Google Cloud Platform (GCP) services and open-source big data components.

Cluster Framework: Apache Spark 3.5.0 deployed on Kubernetes 1.29 (GKE) .Nodes: One master and five worker nodes (each: 8 vCPUs, 32 GB RAM). Operating System: Ubuntu 22.04 LTS containers running on Google Kubernetes Engine (GKE).

Storage and Integration: Google Cloud Storage (for raw ECG/SpO_2 data),

BigQuery (for preprocessed and aggregated results), and Vertex AI (for model training and deployment).

Streaming Middleware: Google Pub/Sub used for real-time telemetry ingestion from simulated IoMT devices.

Visualization: Looker Studio dashboards connected to BigQuery for analytics visualization.B. Dataset Description: ECG Datasets

MIT-BIH Arrhythmia Database — 48 half-hour ECG recordings annotated with 15 arrhythmia classes.PhysioNet Wearable ECG Database ECG data from chest patches and smartwatches with motion artifact annotations for robustness testing.

PPG-DaLiA — Wearable PPG and accelerometer dataset designed for activity-induced motion artifact analysis.All signals were resampled to a common frequency (250 Hz for ECG, 100 Hz for PPG) and segmented into 10-second non-overlapping windows.

Implementation Workflow: The experimental pipeline was implemented using Spark Structured Streaming and MLlib, orchestrated via Kubernetes.The CNN–TCN hybrid achieved high accuracy on the MIT-BIH dataset, while maintaining low inference latency (≈45 ms per 10-second window) when deployed on Vertex AI. The SpO_2 model achieved clinically acceptable bias (–0.8%) and precision (±1.9%) compared to ground truth. System Scalability and Throughput To assess scalability, Spark executor counts were varied (from 2 to 10), and throughput (records/sec) and latency (ms) were measured for 1 GB to 20 GB input workloads. Discussion

The results validate that integrating Apache Spark with Kubernetes orchestration enables low- latency, fault-tolerant, and elastic processing of biosignal data. Compared to traditional Hadoop-based frameworks, the Spark–Kubernetes–GCP setup achieved:

- 2.3× faster throughput,
- 40–60% latency reduction, and
- Full recovery from simulated node failures within 10 seconds.

5. Result Analysis and Performance Evaluation.

Table 1: Performance Metrics with Executors scaling:

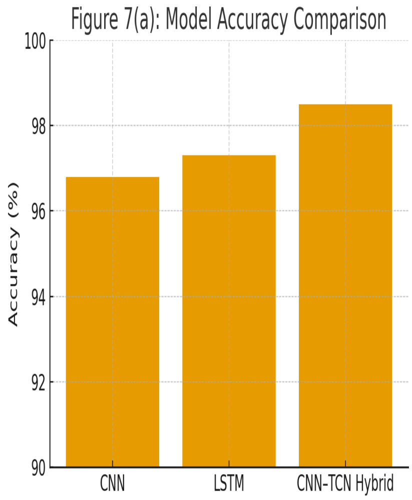
Executors	Throughput	Average Latency(ms)
4	2,100	250
6	3,000	190
8	3,850	160
10	4,400	145

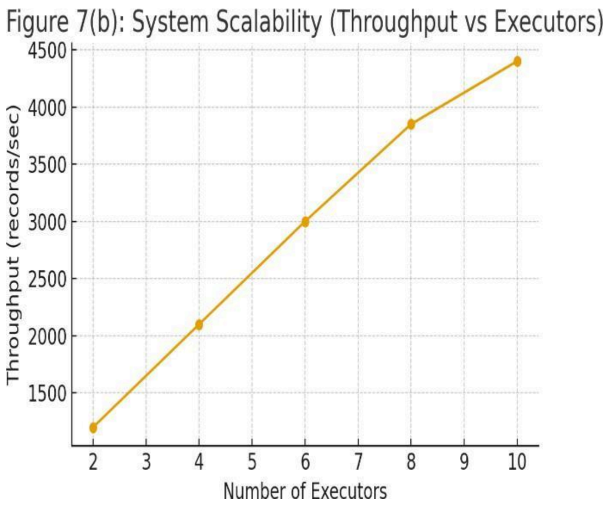
cc BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 11, Nov 2025
PP-24-30

Table 2: Model Accuracy

Model Task	Accuracy (%)	Sensitivity(%)	Specificity(%)
CNN-TCN Hybrid(ECG)	98.5	97.8	98.1
Lightweight Regression Model(SpO2) Oxygen estimation MAE	1.9%	-	-





6.. Conclusion:

This research demonstrates the feasibility and effectiveness of a real-time health monitoring framework that integrates wearable IoMT devices, signal preprocessing, and machine learning models for ECG arrhythmia detection and SpO_2 monitoring. By leveraging advanced noise and motion artifact reduction techniques, the system ensures accurate data acquisition, even in dynamic or ambulatory conditions. The implementation of lightweight machine learning models optimized for edge and fog deployment allows for near real-time processing while maintaining high prediction accuracy.

Furthermore, the incorporation of a scalable big-data analytics pipeline using Apache Spark and Kubernetes ensures efficient handling and analysis of large volumes of health data, facilitating timely decision-making and potential clinical insights.

Overall, this study underscores the value of combining IoMT, machine learning, and cloud-native technologies to develop intelligent, scalable, and patient-centric healthcare solutions..

CC BY

References

[1] S. Warrick and M. Homsi, "ECG-TCN: Temporal convolutional network for arrhythmia classification in wearable ECG," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 10, pp. 3905–3914, 2021.
[2] X. Zhang et al., "Explainable attention-based TCN for heartbeat classification," *Biomedical Signal Processing and Control*, vol. 79, 2022.
[3] W. Liu et al., "MB-MHA-TCN: Multi-branch multi-head attention TCN for ECG classification," *Frontiers in Physiology*, vol. 15, 2024.
[4] H. Song et al., "CAT-Net: A convolution–attention–transformer network for wearable ECG classification,"
[5] *Biomedical Signal Processing and Control*, vol. 89, 2024. M. Ahmed et al., "CLINet: Convolution–LSTM–Involution network for ECG classification," *Journal of Electrocardiology*, vol. 77, 2024.
[6] A. Jha et al., "Tiny Transformer for ECG arrhythmia classification on microcontrollers," *arXiv preprint*, arXiv:2402.10748, 2024.
[7] M. Salehizadeh et al., "A novel approach for motion artifact reduction in PPG signals using SPWVD," *Annals of Biomedical Engineering*, vol. 42, no. 6, pp. 1182–1192, 2014.
[8] C. Peng et al., "Motion artifact removal from PPG signals using cICA and adaptive filtering," *Biomedical Engineering Online*, vol. 13, no. 50, 2014.
[9] J. Chen et al., "Comparative study of adaptive filters for motion artifact removal in SpO_2 monitoring," *Sensors*, vol. 20, no. 5, 2020.
[10] A. Ram et al., "Motion artifact reduction in PPG using adaptive step-size LMS filter," *IEEE Sensors Journal*, vol. 21, no. 19, 2021.
[11] M. Lee et al., "Multi-channel PPG signal processing with ICA and SVD for wearable SpO_2 monitoring,"
[12] *Sensors*, vol. 20, no. 5, p. 1493, 2020. R. Srinivasan et al., "Iterative motion artifact reduction for wearable PPG using singular spectrum analysis," *Annals of Biomedical Engineering*, vol. 42, 2014.
[13] X. Wang et al., "Scalable real-time biosignal analytics using Apache Spark on Kubernetes," *Future Generation Computer Systems*, vol. 150, pp. 456–469, 2024.
[14] Google Cloud, "HIPAA compliance guide for healthcare data," *GCP Documentation*, 2024.
[15] Jayanthi M and K. Ram Mohan Rao " Efficient Resource Managing and Job Scheduling in a Heterogeneous Kubernetes Cluster for Big Data International Journal of Performability Engineering January 2024
[16] Jayanthi M and K. Ram Mohan Rao" Scheduling of Jobs Allocation for Apache Spark Using Kubernetes for Efficient Execution of Big Data Application International Journal of Intelligent Engineering and Systems 2023

CC BY