1 ResuméBusiness Viborg arbejder for at skabe optimale rammer for erhvervslivet i Viborg Kommune og haren ambition om at nå 700 medlemmer i 2025. Medlemsafgang truer imidlertid både organisationensøkonomiske fundament og dens rolle som erhvervspolitisk talerør. For at imødekomme denne ud-fordring er der i dette projekt udviklet en datadrevet prototype, der forudsiger churn og identificerercentrale risikofaktorer på medlemsniveau.Løsningen kombinerer brugervenlig formidling med avanceret maskinlæring i et R-baseret workflow.Seks modeller blev afprøvet, hvor Random Forest blev valgt som slutmodel på baggrund af gennem-sigtighed og forklaringskraft. Feature engineering inddrager bl.a. kontaktfrekvens, eventdeltagelse ogmodtaget erhvervshjælp.Modellen er operationaliseret i et interaktivt dashboard, som understøtter medlemskonsulenternes op-søgendearbejde. ProjekteterudvikletmedrespektforGDPRogdataetikogillustrerer,hvordanenlokalmedlemsorganisation med begrænsede ressourcer kan anvende data strategisk og ansvarligt til at styrkefastholdelsen og engagementet blandt sine medlemmer.2 IndledningBusiness Viborg er en medlemsorganisation, der arbejder målrettet for at skabe optimale vilkår for er-hvervslivetiViborgKommune. Medover600medlemsvirksomhederudgørorganisationenenvæsentligaktør i det lokale erhvervsøkosystem – både som netværksfacilitator, vidensformidler og politisk inter-essevaretager. Ifølge chefkonsulent Michael Freundlich er ambitionen at nå 700 medlemmer og en om-sætning på 2,9 mio. kr. i 2025.Men når virksomheder forlader organisationen, reduceres ikke blot indtægtsgrundlaget – også BusinessViborgs netværkskapital og politiske legitimitet svækkes. Derfor er det afgørende at få indsigt i, hvilkefaktorerderøgerrisikoenforudmeldelse,oghvordanmankanarbejdeproaktivtmedmedlemsfasthold-else.4

Idetteprojektudviklesderendatadrevetløsning,derkombinerertekniskanalysemedbrugervenligind-sigt og som respekterer både juridiske og etiske rammer. Målet er at styrke medlemskonsulenternesbeslutningsgrundlag og understøtte en mere effektiv og målrettet medlemspleje.3 ProblemstillingBusinessViborgerregistreretunderbranchekoden26104793,ogarbejdermålrettetforatskabeoptimalerammerforerhvervslivetiViborgKommune. Somenmedlemsorganisationmedover600virksomhederi ryggen, er relationerne til medlemskredsen helt afgørende, både for at dele viden, styrke netværk ogskabe lokal vækst.I forbindelse med præsentationen af Business Viborg udtalte chefkonsulent MichaelFreundlich: “Vores mål for 2025 er at nå 700 medlemmer og en omsætning på 2,9 mio. kr.”Men når virksomheder melder sig ud, mister Business Viborg ikke kun en indtægt, men også værdifuldeforbindelser, politisk legitimitet og mulighed for at gøre en forskel for erhvervslivet i området. For athandle proaktivt ønsker Business Viborg at få bedre indsigt i, hvad der driver churn og hvem der er irisikozonen.Derforskalderudviklesetdatadrevetværktøj,somkombinerertekniskanalysemedbrugervenligindsigt.Et værktøj, der gør det muligt for både medlemskonsulenter og ledelse at træffe kloge beslutninger oghandle i tide med respekt for både dataetik og jura.4 ProblemformuleringHvordankanBusinessViborganalysereoganvendemedlemsdatatilatudvikleetbeslutningsunderstøt-tendedashboard,derforudsigerchurnogforklarercentralerisikofaktorer–baseretpårelevantemaskin-læringsmetoder og med inddragelse af etiske og juridiske overvejelser?4.1 UnderspørgsmålEksplorativ analyse (EDA)Beskriv hvilke mønstre og karakteristika kendetegner de virksomheder, der forlader Business Viborg?5

Modelvalg og performanceHvordan kan forskellige machine learning-modeller anvendes til at forudsige churn i Business Viborgskontekst, og hvilke modeller er mest velegnede?DatavisualiseringHvordan kan resultater og churn-indsigter formidles via et brugervenligt dashboard, som understøtterdaglig opsøgende indsats for medlemskonsulenter og ledelse?Etik og juraHvilke juridiske krav (fx GDPR) og etiske overvejelser bør indgå i udviklingen og brugen af et churn-forudsigelsesværktøj baseret på medlemsdata?5 AfgrænsningI udviklingen af en datadrevet churn-model for Business Viborg er det nødvendigt at foretage en rækkemetodiske og praktiske afgrænsninger for at sikre projektets gennemførlighed og fokus. Følgende un-derafsnit præciserer, hvordan projektets omfang er afgrænset i forhold til teknologisk anvendelse, data-grundlag, modeller, systemintegration og juridiske vurderinger.5.1 AI-chatbots og anvendelse af ChatGPTChatGPT 4.0 har været anvendt som et understøttende værktøj i forbindelse med idéudvikling, sprogligformulering og grammatisk korrektur. Modellen har alene fungeret som et supplement i arbejdet medtekstbaserede opgaver og har ikke erstattet selvstændig analyse, faglig vurdering eller besvarelse af pro-jektets problemformulering. Chatbotten er således ikke anvendt til at generere indhold i den analytiskeeller metodiske del af projektet.6

5.2 DatagrundlagProjektetbaserersigudelukkendepådetdatasæt,dererstillettilrådighedafBusinessViborg. Datasættetindeholder oplysninger om medlemskab, virksomhedsdemografi, branchetilknytning, kontaktaktivitet,eventdeltagelse samt ydet rådgivning. Alle data er pseudonymiserede og begrænset til et afgrænset tid-srum. Dette kan påvirke modellens generaliserbarhed over tid og dens evne til at indfange nyere ten-denser i medlemsadfærd.5.3 Modellens omfang og valg af algoritmerFormålet med projektet er at udvikle en forklarlig og anvendelig prototype frem for en produktionsklarløsning. Dererderforikkeforetagetomfattendehyperparameter-tuningforallemodeller. Seksmodellerertestet–herunderSupportVectorMachine,RandomForestogXGBoost–ogperformanceerevalueretpå baggrund af F1-score og AUC som de primære metrikker. Fokus har været på at finde en balancemellem prædiktiv nøjagtighed og forklaringskraft.5.4 SystemintegrationDenudvikledeløsningerimplementeretsomenwebbaseretprototypeiRogerikkeintegreretmedBusi-ness Viborgs interne systemer, såsom CRM- eller medlemsdatabaser. Modellen kan tilgås og anvendeslokalt gennem RStudio Cloud eller ved afvikling på en dedikeret server, men kræver manuel opdateringaf data. Fremtidig integration og automatisering er oplagte skridt i en potentiel videreudvikling.5.5 Juridiske og etiske vurderingerProjektet indeholder en overordnet vurdering af de juridiske og etiske rammer med fokus på datamin-imering, transparens og behandlingsgrundlag i henhold til GDPR. Der er ikke foretaget en fuld juridiskgennemgang, og tekniske løsninger som adgangsstyring, kryptering og samtykkehåndtering er ikke im-plementeret i prototypen. Disse aspekter betragtes som en integreret del af en eventuel implementer-ingsfase og bør afklares i samarbejde med relevante juridiske rådgivere og systemansvarlige.7

6 DefinitionerI dette afsnit defineres centrale begreber og forkortelser anvendt gennem rapporten:Churn:Når en virksomhed ophører med sit medlemskab i Business Viborg. I datasættet angives dettesom en binær variabel, hvor 1 betyder churn og 0 betyder fortsat medlemskab.Churn-model:Enprædiktivmodel,derestimerersandsynlighedenfor,atenvirksomhedchurner. Dener baseret på historiske medlemsdata og konstruerede forklaringsvariable.Feature Engineering:Fremstillingafnyeforklarendevariablefraeksisterendedata,somstyrkermod-ellens evne til at forudsige churn. Eksempler inkluderer medlemsanciennitet, kontaktaktivitet og delt-agelse i arrangementer.MeetingLength:Længden af det seneste dokumenterede møde med en virksomhed, målt i minutter.Bruges som indikator for relationens styrke.har\_haft\_kontakt:En binær indikator for, om virksomheden har haft kontakt med Business Viborg(f.eks. møder, telefonopkald eller rådgivning).deltaget\_i\_event:Binær variabel der angiver, om virksomheden har deltaget i mindst ét event i anal-yseperioden.hjælp\_kategori:En kategorisk variabel, der angiver typen af erhvervsfaglig støtte virksomheden harmodtaget. Kategorierne er fx Strategi Udvikling, Organisation og Ledelse, Jura og Struktur m.fl.medlem\_antal\_år:Antal år virksomheden har været medlem, beregnet som forskellen mellem anal-ysedato og oprettelsesdato.Machine Learning (ML):En metode til at bygge modeller, der kan lære mønstre i data og forudsigefremtidige hændelser. I projektet er ML anvendt til churn-forudsigelse.Random Forest:En ML-algoritme, der kombinerer mange beslutningstræer for at skabe en robust ogforklarlig model. Valgt som slutmodel i projektet.ROC AUC:Etmålformodellensevnetilatadskillechurnereogikke-churnere. Enværditætpå1indik-erer høj prædiktiv nøjagtighed.8

F\_meas (F1-score):Et samlet præstationsmål, som balancerer præcision og recall – særligt velegnetved skæve datasæt.Dashboard:Et interaktivt visualiseringsværktøj, der præsenterer churn-risici og medlemsindsigter påen overskuelig måde til brug i den daglige medlemspleje.GDPR:EU’s databeskyttelsesforordning. Projektet tager højde for centrale principper som datamin-imering, transparens og legitimt behandlingsgrundlag.7 Analyse7.1 Dataforståelse og fordelingAnalysen tager udgangspunkt i et datasæt bestående af 2.966 medlemsvirksomheder tilknyttet BusinessViborg, som har et selvstændigt P-nummer. Datasættet afspejler en betydelig variation med hensyn tilvirksomhedsstørrelse, branchetilhørsforhold og interaktionsniveau med organisationen.Enindledendefordelingafslører,atvirksomhederudendokumenteretkontaktellerdeltagelseiarrange-menterharmarkanthøjerechurn-rate. Denneobservationantyder,atfraværetafkontaktogengagementkan være centrale indikatorer for medlemsophør.7.2 Egenskaber ved virksomheder med høj churnForbedreatforståhvordanforskelligeformerforengagementpåvirkermedlemsstatus,harvikombineretto centrale variabler: kontakt med Business Viborg og deltagelse i events. Dette giver fire grupper, somvarierer i deres relation til organisationen. Figuren nedenfor viser tydeligt, hvordan virksomheder medbåde kontakt og eventdeltagelse i langt højere grad fastholdes som medlemmer, mens fravær af beggefaktorer er stærkt forbundet med churn.9

Figure 1:Fravær af begge faktorer er tæt forbundet med udmeldelse, mens dobbelt engagement viserstærk fastholdelse.Virksomheder med begrænset kontakt til organisationen, lav deltagelse i arrangementer og uden doku-menteret interaktion har generelt en markant højere risiko for at opsige deres medlemskab. Dette un-derbyggesaffigurennedenfor,derviserdefempostnumremeddenhøjestegennemsnitligechurn-risiko.Herses,atgeografiskeområdermedlavtilknytningtildetcentraleområde(8800Viborg)udvisersærlighøj churn-sandsynlighed.10